From communities to crowds: quantifying the subjective

Data, measurements and models

Sagar Joglekar

Department of Informatics, Faculty of Natural and Mathematical Sciences

King's College London

This dissertation is submitted for the degree of

Doctor of Philosophy

February 2020

– To Geetika, Medha , Chanda ... and now Arya.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. The dissertation covers contributions for journals and conferences where I was the main contributor and primary investigator.

Sagar Joglekar
February 2020

ACKNOWLEDGEMENTS

Behind this dissertation, is a battalion, who either inspired, nudged, guided, mentored or supported me, to eventually pursue something which had been a dream in making for almost 20 years.

Firstly, I have to express my deepest gratitude to my advisor and mentor Dr. Nishanth Sastry. I still remember the long and detailed email he wrote, even before I joined his group, unwrapping every doubt I had about all things Ph.D. It was because of this email and the brutally honest but empathic advice within, is why I decided to take a leap of faith and pursue this journey under his wing. At every step, he has inspired me to explore the difficult and to be fearless doing so. With a deliberate balance of control and freedom, his style of mentoring is something I am truly thankful for and wish to internalize at some point.

The biggest curse of doctoral studies is the loneliness that it breeds. But I was immensely lucky to find friends on and off campus. I am truly grateful to find friends like Aravindh Raman, Changtao Zhong, Dmytro Karamshuk, Pushkal Agarwal, Emeka Obiedu and Miriam Redi, with whom I have had the pleasure of sharing countless coffees and endless inspiring chats. Some of these chats have even blossomed into full fledged research projects and generated grants and/or publications. I have been blessed with friends like Anand, Sanika, Marc and Roxy, who filled our life with fun, love and a warm feeling of being part of a family.

I have been fortunate in life to find mentors at different stages, who have shaped my career and life in some way. Here too, I was lucky to find mentors either through collaborations or shared interests, who have been quintessential in shaping my Ph.D. and my career. I would like to thank Luca Aiello, Miriam Redi, Gareth Tyson, Daniele Quercia, Pietro Panzarasa and Anna Simoni for their support and valuable mentoring.

Finally, it only makes sense to connect the dots in hindsight. To that end, I have to thank my friends and colleagues from Santa Barbara who gave me honest and balanced advice when I needed the most. I would like to thank Advait, Geetak, Bugra and Gina, who made sure my desire to do a Ph.D. became a drive and resulted in an action.

My decision to pursue a Ph.D., could not have been possible without the unconditional support of my family. Every curious thought in my mind can be traced back to how I was raised. My mother Medha and my sister Chanda, have been the biggest driving force behind all my actions. They have always inspired me to follow my heart, even if at times it meant me going away from them for a long haul. I would like to thank my other half of the family, Sanjeev, Sandhya and Rishikesh, who had faith in me when I needed it the most.

Last, but by long shot not the least, I have to thank my wife Geetika, who has been like a light house in a turbulent sea. Unwavering in her support and unparalleled in her guidance, she has gone above and beyond, to make this dissertation a reality. And hence I believe it is as much hers, as mine. And to conclude, I would like to thank Arya, our new born daughter. She made sure that I kept to the schedule of this Ph.D., by announcing hers when I started writing up.

ABSTRACT

Research in the field of recommendation systems have shown that *subjective preferences tend to follow deterministic patterns, when looked at data in large sample sizes*. This principle underpins several of our present day online and off-line recommendation applications like e-commerce , restaurant recommendation, place recommendation or music recommendations. With the ever pervasive nature of the internet, we as a society have gone beyond treating the online spaces as a tool to access information, and have started treating them as a natural extension of the self. We spend more time than before as a part of a larger networked community, exchanging thoughts, debating ideas, expressing creativity and "socializing". We also sometimes indulge in expression of human emotions like empathy, anger, sadness and sometimes seeking help. On the other hand, at times we behave like crowds; participating in entertainment and engagement channels and offering a piece of our attention budgets, without the explicit intent of being social. But in both cases, our decisions are often governed by our subjective perceptions of the online and offline worlds. At such a juncture, I examine the central thesis **Can we quantify properties of subjective nature, if the data is large enough, and originates from human communities or crowds?**. In this dissertation I develop data driven methods with the aim to quantify subjective qualities, through two case studies. I investigate the utility of said methods in designing interventions to improve the online and offline spaces. I do so by testing the validity of the central thesis in the context of two distinct scenarios, the first dealing with quantification of social support by looking at the network structure of online support networks and the second with quantification of subjective properties from crowd sourced opinions.

In the first study, I analyse on-line spaces involving networked communities where interactions between humans are purely with the intent of helping each other. In this study, I test the I develop frameworks to abstract out the graphical structure of these interactions. Using these abstractions, I examine these support communities from a macro scale to understand the signature behavioural patterns that makes these communities thrive. I then investigate presence of perceived support by finding discriminative local and global structures in these communities. I argue that these structures, which we call anchored motifs, are the

signatures of a supportive exchange process in online conversations. This informs my analysis about the nature of peer support in these communities and paves the way to do actionable interventions in the area of perceived support in online networks.

In the second study, I investigate utility of crowd's perception of aesthetics in physical spaces. As such the aim is to explore the potential of crowd perception in developing tools, to better design physical spaces. I do this by developing a pipeline that capitalizes on crowd sourced responses about perception of urban aesthetics. I develop a deep-learning driven framework, which is able to quantify the perception of intangible qualities like 'beauty of a space' through a crowd sourced rating of google street view images. I show that a general pattern of beauty in urban spaces can be learnt through crowd sourced opinion and deep learning models. I further develop a generative model to simulate beautification of urban spaces. Through a detailed literature review of the field of urban design, I develop a measurement framework which can provide insights into the predictors of urban beauty. I then develop the necessary tools to evaluate these metrics using computer vision techniques. I validate the value of these metrics through an expert survey and also validate the interventions using crowd sourced perception experiments.

Above all, in this dissertation, I contribute original frameworks and implementations for different approaches towards quantifying subjective signals from communities and crowds. These methods verify and validate several metrics developed for understanding subjective properties, like perceived support and perceived aesthetics, at scale. This provides a path forwards for A.I. driven design and curation of online and offline spaces.

⌈ ⌉

TABLE OF CONTENTS

LIST OF FIGURES

List of tables

CHAPTER 1

INTRODUCTION

There are things known and there are things unknown, and in between are the doors of perception - Aldous Huxley - *Doors of perception*

1.1 Introduction

We live in a world where information is being bombarded on our cognitive faculties from all sides, at all times. The internet is a continuous stream of information and each source is fighting with the other to get a piece of our attention budget. With the advent of machine learning and big-data, building systems that predict actions as a response to perceptual triggers has become the bread and butter of many companies. The use cases may range from understanding which adverts made a visitor do an unscheduled purchase, or which string of music tracks recommendations maximized a users time on a particular music platform. But in the end it all boils down to understanding what triggers result in human action or lack there of [SDP12]. Nevertheless, the systems that surrounds a human interacting with the internet are all designed to figure out the best triggers which are perceived by the human as worthy of attention. Since the dawn of data driven design, the user's attention has become such a valuable commodity, such that the academic community has coining a term for it – "The Attention Economy" [DB01]. In the words of Matthew Crawford "*Attention is a resource, a person has only so much of it.*" [Cra15]. And this very resource has been the driving force behind the advertisement driven internet giants like Facebook and Google.

We live in the age of distraction, and more often than not our subjective perceptions are guiding our actions, instead of our conscious cognitive processes. Several studies have shown that engagement is almost always a game of stimulating our most basic urges, such as dopamine hits, presence of faces or simply arousal of emotions to increase the working memory [BSG14, JSR17, SFSJ06, Soa15].

On the positive side of this story, the capacity of subjective perceptions to influence actions, is also being used to design positive interventions. One such example is the emergence of more formal topical spaces on the internet, that facilitate providing peer-to-peer perceived social support [Cou05, BSBMDS16] or information exchange [FM08]. The ever pervasive nature of the internet allow these formal spaces to function almost like physical communities, with moderated and effective peer to peer exchange of thoughts, ideas and empathy [KGB$^+$02, Squ15, HOG$^+$10].

There has also been a series of works [JK19, QOC14, QSAM15, Que15, QSA14] that looked into whether we can quantify the perception of the physical world based on the crowd-like behaviour of individuals online. Here crowd-like implies that there is no social interaction between two individuals generating the interaction data. These individuals take actions online(check-in on a platform, post a picture, tweet etc) influenced by their perceptions of the real physical space. In these situations, their perception of real spaces is influencing what they post, like, do online.

Both these examples show that in the age of the internet, our online and offline lives have been linked more deeply than ever. Our once assumed offline personal needs, like social support, are being addressed via online forums. At the same time, actions that we take offline are enriching and influencing our online presence. In both cases the subjective perception of our online and offline environment is impacting or life experience. This provokes the question:

> Can quantifying properties of subjective nature help us design impactful interventions for our on-line and offline lives?

This question has been the motivation behind all the works done in the past 4 years, eventually resulting into the formation of the central thesis of this dissertation. In the effort to pursue this question, we first need to understand what does it mean to "quantify the subjective", in the context of web-scale data. But first, we need to clarify the framework that grounds this dissertation's approach towards perception, affect and data. To do so we should try and understand each of these terms separately.

1.2 Perception and Affect

In this dissertation, I tried to build frameworks to capture the signatures of perception of the subjective. This was accomplished using large volumes of data and metrics designed around concepts from inter disciplinary fields. The utility of such an attempt, can only be justified if there is a real link between how humans take decisions at the most fundamental cognitive level and how they perceive the world around them. If there exists such a link, then the signatures found in the data can be explained and capitalized.

There has been an ongoing effort to unravel this link, through psychological, neuro-evolutional and philosophical arguments. I will try to gain inspiration from them, but a detailed critique is beyond the scope of my dissertation and expertise

Definition 1 ***Affect*** [1]*: Any experience of feeling or emotion, ranging from suffering to elation, from the simplest to the most complex sensations of feeling, and from the most normal to the most pathological emotional reactions.*

Definition 2 ***Perception*** [2]*: The process or result of becoming aware of objects, relationships, and events by means of the senses, which includes such activities as recognizing, observing, and discriminating. These activities enable organisms to organize and interpret the stimuli received into meaningful knowledge and to act in a coordinated manner.*

[1] American Psychological Association definition.
[2] American Psychological Association definition.

Emotions or ‘affects’ and perceptions have long been discussed in the psychology, neuroscience and philosophical literature. The two are tightly linked in our understanding of mind. Perception is a process, and affect is a result that this process sometimes evokes. Emanuel Kant in his prolific work, first discussed the utility and the philosophical reasoning behind presence of affects or emotions [Kan87]. In his opinion, emotions are pre-cognitive involuntary states, termed as "mere perceptions of unspecified bodily states" [Bor04]. But according to him, that does not mean they don’t influence our deepest level of well-being and decision making processes. The link between affect and perception has also been explored in several other cases. One way to look at it was through the framework of understanding the aesthetic. In some sense, the aesthetic is an affect inducing entity. The act of enjoying a gorgeous sunset or a beautiful flower defines a sentient perceptive mind, as much as language or art. An argument to link the aesthetic with perceptions was made by Perlovsky [Per14], where they propose that the phenomenon of affects arousing from perception of aesthetics, comes from a fundamental human need to enrich the knowledge about real world. An unexpected thing, stimuli or structure in physical space creates a dissonance between our expected model of the world and the perceived reality. And at some level we perceive it as aesthetically pleasing. Another recent study by Zadra et.al [ZC11] evaluated the relation between visual perception and emotions. They demonstrate that the conventional assumption of the disentangled functioning of perception and affects is not true. Humans are quite susceptible to perceiving different realities based on different aroused affects. So a happy person would perceive a 50$ shoe to be a reasonably priced item, which a sad person may not.

The discussion on the formal definition and process of affects will continue, but there seems to be a consensus, at-least among the computer science and information science community that affects do influence our decisions and we perceive information through a filter of affects. Affective triggers can be generated when information is formatted or packaged in a certain way.

In such a setup, it is worth testing if certain affect driven interactions on the web leave a trail of patterns in the data of these interactions. Furthermore it is worth asking if these

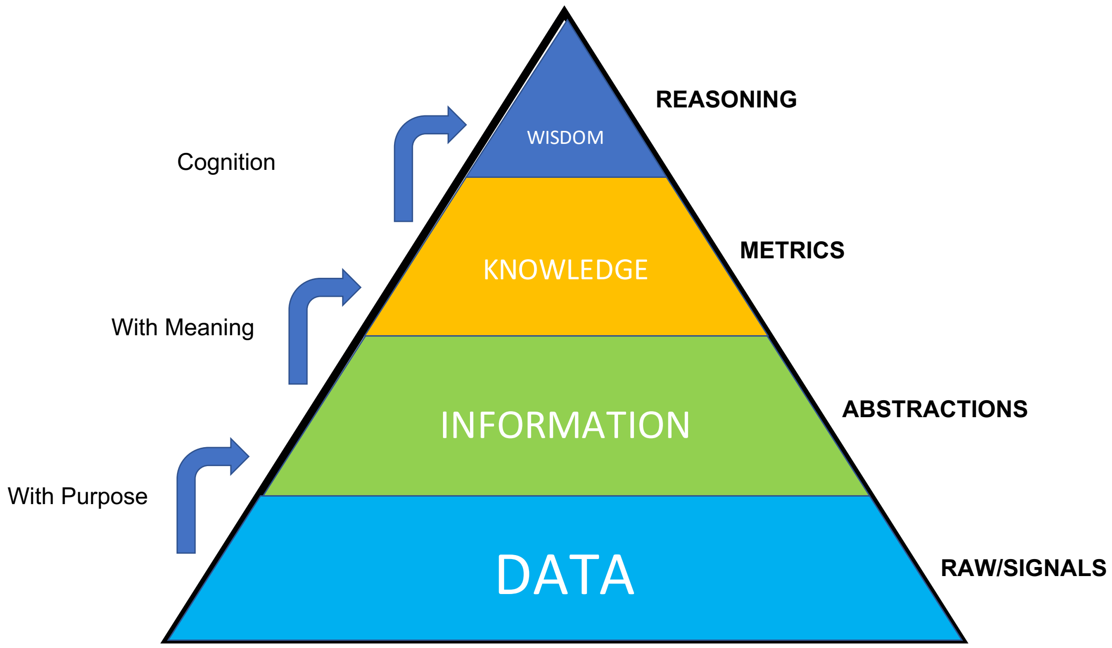


Fig. 1.1 The DIKW pyramid: The four layers of the pyramid represent the four distinct stages that the data goes through. The terms on the left hand side of the pyramid indicate the intent which results into the transformation of one layer into another. The terms on the right indicate the outcomes of each transformation.

patterns might in some way be used to improve, through interventions, our online and physical environments. But to arrive at these patterns, one needs to formalize the frameworks for approaching such a problem. The journey from data to the subjective signatures needs to be standardised.

1.3 The DIKW model

My work in this dissertation involves building pipelines that could extract signals about human processes from large scale data. But to make the process of building pipelines repetitive across use cases, we need to first settle on a framework that enumerates our intentions about the data. On the journey to develop the pipelines, one has to reflect on frameworks that codify operations on data. One such framework, which I found most relevant to my work, is the **Data , Information, Knowledge and Wisdom** model [Row07].

The original framework was theorized to elaborate the journey between raw data and actionable wisdom. To that end I link these pyramidal layers with the different outcomes of the pipelines I developed during this dissertation.

With the pyramidal structure of this framework, with each layer the data gains more structure and the relationships become more defined.

In this model, the most foundational layer consists of the pure form raw **Data or signals** that come from a source. If we are measuring subjective perceptions of humans, this source needs to be tied back to humans in some way. To that extent, the data must ideally be a product of human to human interaction online. Or it needs to capture the human perceived responses, through explicit exercises like crowd sourcing or public surveys. In this dissertation, I present solutions in the context of exactly these use cases, where the data sources are either from human to human interactions or a result of crowd based perception.

The **Information** layer compels any processing done on the data, to be with a sense of purpose or an end goal. For example, if the goal is to understand how humans exchange messages at times of distress, you would most certainly need to express the raw information about sender and recipient of messages into some form of a networked abstraction. This compulsion of goal generally forces us to choose an abstraction to structure the data into. The abstraction preserves the organization of data, but at the same time allows information to be operated on.

Knowledge. Defining knowledge has been an ongoing effort in the field of philosophy. But in the context of information science, knowledge involves collation of diverse sources of information and mix of contextual information, values and metrics to deliver a coherent understanding of the real world. For example, if you need to know the most popular user among a social network of users exchanging messages; you would look for the most central user in the network(abstraction) along with several other temporal and structural metrics to arrive at a few candidates. In this particular case, these metrics, along with the context of the social network's design, dawns the meaning of "popularity".

The final layer requires a literature driven approach that links the outcomes of the previous layer (the metrics), with a defined set of concepts and vocabulary in the field of intervention.

The very act of gaining insights using these metrics is what we can call **wisdom**. In order to maximize impact of this process, we need to make sure that we are able to reliably map the metrics that we design onto a set of literature backed vocabulary in the field of practitioners.

In our example, let's assume we need to get some insights about the dynamics of popular users. Particularly in the context of optimizing advertising delivery. For example we need to understand how a particular piece of advertising might percolate through the network if certain popular users advertise it [LS12]. However , to arrive at these insights we need to be grounded in the vocabulary of epidemiology, network physics and depending on the application, advertising or meme theory. Then using the abstractions of social networks and the metrics derived from them, one can design a pipeline that maps the metrics onto concepts like "virality" , "influence" or "contagion".

This pyramidal approach inspires all the frameworks and data processing pipelines I developed in this dissertation. Figure 1.1 shows an illustration of the adopted version of Rowley's DIKW model, which I would refer back as a repeating motif throughout my dissertation.

1.3.1 Data

Data is one of the most fundamental contribution of this dissertation. To develop frameworks around quantification of human perceptions, so that we can do impactful interventions, we first need to make sure we formalize how we acquire, clean and condition our data. The foundational level of this pyramid is the data that the framework would work with, in order to ascend towards extracting wisdom. I worked with diverse forms of data such as textual data , video data and image data to understand how these might exhibit signatures of human perceptive processes. The relation between data and subjective attributes needs to be examined using some proxy such as interactions on a post, votes or engagement. For this reason, my research involved collecting data from sources where either human to human interactions happen, or the data is generated on account of a human expressing their opinion about a subjective quality, like in case of crowd based perceptions.

Interaction Data

The first case study of this dissertation focuses on online support communities, where human to human interactions which express support, are at the centre. It has been shown through several studies in medical informatics, that these communities play a very important role in providing support and respite in times of distress [AVKR16, MC12, PS15, BC11, IABS^{+}17, HBCF16]. The communities are especially helpful when it comes to people suffering from long term illnesses or mental health issues. The key element that impacts the users is the perceived social support [Nam11], which delivers people in distress a sense of belonging to a group and a sense of empathy from the fellow supporters. To understand how users on these communities perceive social support, I work with data acquired from online health forums, where users share, give support and ask for support. I look at communities that deal with long term conditions like Lung illnesses, and communities where mental health patients seek support [JSC^{+}18]. The data spans across a duration of 10 years, containing peer to peer support interactions of more than 30,000 users. I also crawled a popular forum based social network called Reddit [3] to acquire a peer to peer support data regarding mental illness and suicidal thoughts. The data covers discussions around more than 30,000 calls for support, and includes the complete structure of the way people respond to these calls.

Image data

The other facet of my work looks at the problem of quantifying our perception of physical spaces. Whether a street is considered beautiful is a matter of subjective opinion. And yet research has shown that there are specific urban elements that are universally considered beautiful: from greenery, to small streets, to memorable spaces [Ale77, QOC14, SSH13]. These elements are those that contribute to the creation of what the urban sociologist Jane Jacobs called 'urban vitality' [Jac61]. Apart from vitality, these motifs in urban environments are also highly correlated with feeling of well-being, health and safety [KK89]. There have been studies where people have tried to use crowd sourcing to acquire subjective ratings of

[3] www.reddit.com

images [SPM15] which have shown some reasonable progress on this front. But the real gap in these studies is understanding the impact of urban elements on the perception of these subjective qualities. E.g. How much does presence of a green garden affect the subjective rating of beauty for that particular area? For this reason, I work with google street view data and crowd sourced subjective ratings of various places around the world [NPRH14], with the aim to understand how people perceive the sense of beauty in urban areas. Then using the semantics of urban design and architecture, developed by a detailed literature review, I aim to develop machine learning pipelines that can suggest interventions to change perceptions of physical spaces.

1.3.2 Abstractions

The act of aggregating information from data, almost always involves building structured abstractions. Throughout my dissertation, I either repurpose well known abstractions in computer science or develop my own using tools from fields like computer vision and Information theory. For the first study, I incorporate user meta data and the textual data of their activity, to build organized networked abstractions representing the conversation structures on the support forums. I use these abstractions to evaluate global and local structures in support communities, which would be discussed in detail in Chapter 3 and 4.

While working with image data, I use several pixel level abstractions to segment and group semantically similar pixels. I also use several state of the art object and scene detection algorithms, to extract semantic information from an image, with the aim at analysing correlations with the perception of subjective attributes of images with these metrics. I also use deep convolutional networks and generative models, to abstract out a representation of beauty. There will be a more detailed discussion of these abstractions in the later chapters (Chapter 5 , Chapter 6).

1.3.3 Knowledge

For extracting knowledge, we need to first associate meanings to certain computable metrics that we obtain from the abstractions. As discussed in the previous example, it could be as simple as associating the property of "popularity" to the metric of centrality. In this work, I developed several such metrics which related subjective properties with measurable structures in data. These structures could be in the form of topological motifs in the interaction graphs, or in the form of patterns in images. Some of these metrics are based on intuitions which I validate, and some based on extensive literature survey. To give an example, I develop the concept of anchored triads, which combines local structures in interaction graphs of users, with the positional parity of a user in the conversation. This helps me understand how a group of people behave around a person in distress in a supportive conversation, as against a general conversation.

1.3.4 Wisdom

Finally the wisdom underlies insights that come for experience. The experience could come simply from the scale of data or from cross disciplinary literature that puts forth theories of subjective experience. E.g. The theory of social support puts forth four categories of social support 1)Affective/Perceived 2)Instrumental 3)Informational 4)Networked [CS92]. Each type has its own specific traits. My dissertation looks at these theories from the lens of computational social science and networks science to develop metrics and methods to partly quantify signatures social support.

1.4 Research Thesis and Research Questions

To make any progress towards answering the guiding principle as articulated in 1.1, we first need to make progress towards quantifying the subjective. Hence the overarching central thesis of interest that I would explore through the two case studies is:

Can we quantify properties of subjective nature, if the data is large enough, and originates from human communities or crowds?

But this thesis question is quite open ended, and answering it in a generalized manner seems impractical in the scope of one Ph.D. For this reason, I need to first contextualize my work in the realm of practical applications, by deriving more focussed research questions, such that I can acquire data and test my hypothesis in an effective time bound manner. More so, to drive maximum impact, I would like to focus on applications which have the potential to have real world consequences, either through interventions or through inspired interest in the field.

1.4.1 PART 1 : Interactions on support communities

Humans are social animals in every aspect. The presence of social support systems in ones lives have shown to have huge quantifiable benefits. From speeding up recovery in cases of post-partum depression or in the cases of cancer survivors [CDSLS93, DS84, BCH$^+$90] , to signs of positive turn around among patients suffering from alcoholism and depression [PFR$^+$00, BAH$^+$86], social support is a key predictor of positive prognosis for patients under distress. With the advent of internet, a lot of communities have sprung up, which provide a rich platform for patients to interact, exchange support as well as provide a perceived sense of community. These communities are moderated, only to an extent to curb toxic behaviour, but other than that are largely free form. Due to a very homogenous membership, where most members have either gone through or are going though similar distress, there is an emergent sense of support and affective empathy [DSSBSM16]. The key idea of this case study was **to quantify the signatures of social support in the behaviour and interaction patterns on these supportive communities**

Social support, or perception of help received from others, is a widely studied as a psychological resource used to cope with distress. The social support is generally classified into one of the following categories viz. Informational, Tangible , Network or Emotional(affective) [CS92]. These categories measure the nature of social support along the idea of exchange of resources between the person in distress and the one providing support. For example, an informational support interaction could involve pure exchange of valuable information about dealing with an issue. Whereas as network support interaction could purely allow the recipient to acquire a wider network of people. In the part 1 of my thesis, I would like to quantify the nature of perceived social support from the structure of the dialogue and interactions between people. I conjecture that in doing so, we can tease out the signatures of network and informational support in online communities.

Despite the highly specific nature of support communities, they have been succesful in maintaining their relevance. In such a context, it is highly valuable to examine the dynamics communities and understand the macro level behaviours of the users. Moreover, it is worth understanding how the interactions on these communities differ from generic interactions on the web. The discriminative differences are what I call signatures of support. Quantifying these signatures would facilitate these communities to be better poised to tackle any disturbances in the dynamics of social support. These signatures would also help us quantify the net utility these communities in terms of delivered informational or network support. With this in mind, I formulate the following research question for my first case study.

RQ1 *What dynamics of support communities help them thrive?*

RQ2 *What differentiates **users** on support communities from generic ones?*

RQ3 *What differentiates **interactions** on support communities from generic ones?*

In order to make progress on these questions, I first had to collect data from two different communities designed for online social support. The first community is dedicated for patients suffering from chronic lung diseases, such as Asthma or Chronic Obstructive Pulmonary Disorder (COPD). This community was moderated by self appointed moderators, and everyone on this community was either a survivor or a patient of these diseases. This community allowed patients to ask questions about symptoms and home remedies and sometimes just bond over social interactions. The second community I worked with dealt with people suffering from chronic depression and suicidal thoughts. This community was a safe haven for such people to vent out suicidal thoughts and get support from peers to manage these sudden flares of thoughts of self harm.

Through these two communities, I develop a pipeline to analyse the peer to peer interactions using abstractions derived from network science. The abstractions capture conversation level structure as well as the global interactions in a community. This allows me to probe the evolution of such a community both in terms of overall structure of the network as well as conversational interactions between users. I also develop metrics inspired from psychology and psychotherapy literature to quantify how these interactions can be qualified as supportive or non supportive. For example, the concept of communication accommodation theory (CAT) [CG88] prescribes the way in which three participants can participate in a group therapy session. The metrics that measure the branching of a conversation or the patterns in ties between 3 users (triads) align with these prescriptions.

Through a data driven analysis, I establish confidence on these metrics. Through this process, I also report my findings about the dynamics of users on these communities and key properties of user roles. I find that these conversations have a distinct nature when compared against regular baseline conversations over the web, and these distinct signatures could one day be used to curb toxicity as well as improve the support community interface.

1.4.2 PART 2 : Leveraging crowd's perceptions about aesthetics of real spaces

Urban aesthetics and presence of certain elements in the physical spaces that we use, have shown to have lasting effects on our mental health[SPM17] and physical well being[BBLO01, GCBK$^+$05]. However, with the advent of access to large scale data and machine learning techniques, we have a unique opportunity to quantify what exactly comprises of urban aesthetics. In the next part of my dissertation, I aim at using the scale of the internet to try and improve how our cities are perceived. In part 1, the interactions between networked users helped me understand the signatures of social support. In this part, we investigate if the opinions of a large number of disconnected users (crowds) can help us quantify something as subjective as the sense of beauty. In this study, I investigate the following research question:

RQ4 *Can opinions of disconnected crowds help us model the perception of aesthetics in real world?*

RQ5 *Can machine learning models based on crowd opinions help us improve real spaces?*

RQ6 *How much do the suggested improvements align with the expectations from the literature or the practitioners?*

Crowdsourcing is a method through which one could get inputs, subjective or otherwise, about a particular set of questions from a large number of real humans using the internet. In return the participants could be offered a tangible compensation, or in some cases, a gamified incentive. The **RQ4** motivates me to investigate if we can use crowdsourcing to quantify how people perceive urban spaces. Research has shown that if a large number of people could vote on a set of images, regarding their aesthetic quality, a trend emerges that favours some objective metrics of beauty[D$^+$08, QOC14]. Can we link these metrics to urban elements? For this reason I work with google streetview images, where real people vote on aesthetic value of images through a large scale crowdsourced study. After evaluating for statistical trends in preference of aesthetic urban images amongst the voters, I attempt to answer **RQ5** by training a deep convolutional neural network model, which can discern between an aesthetically pleasing and unpleasant urban scene with a high degree of accuracy. Once we have a model that could "detect" beauty in urban scenes, we could then use machine learning and deep learning techniques to understand how different urban elements relate to the notion of beauty. I further develop a set of metrics which can explain the differences between a beautiful and ugly image in the vocabulary of urban design. These differences are suggested as hints to practitioners, in order to improve an existing urban space. These hints are given in the form of suggested changes in different popular urban design metrics, which makes the whole process legible to practitioners in the field. I further test the validity of these metrics, by comparing them against expectations of the literature and those of real urban design practitioners (**RQ6**).

1.5 List of peer reviewed publications

I would like to list all the publications which resulted from the past 4 years of work, as well as collaborations I was able to strike with a diverse group of researchers. The author lead publications have influenced different chapters of this dissertation.

1.5.1 Original author contributions

List of papers (published, accepted and under peer review) which were either led by the author or where the author had a fundamental contribution

1. **Joglekar S**, Sastry N, Redi M. Like at first sight: understanding user engagement with the world of microvideos. In International Conference on Social Informatics 2017 Sep 13 (pp. 237-256). Springer, Cham.

2. **Joglekar S**, Sastry N, Coulson NS, Taylor SJ, Patel A, Duschinsky R, Anand A, Evans MJ, Griffiths CJ, Sheikh A, Panzarasa P. How online communities of people with long-term conditions function and evolve: Network analysis of the structure and dynamics of the asthma UK and British lung foundation online communities. Journal of medical Internet research. 2018;20(7):e238.

3. **Joglekar S**, Velupillai S, Dutta R , Sastry N "Online discussions about mental health in Reddit exhibit signatures of supportive conversations" Under Review

4. **Joglekar S**, Quercia D, Redi M, Aiello LM, Kauer T, Sastry N. FaceLift: a transparent deep learning framework to beautify urban scenes. Royal Society Open Science. 2020 Jan 16;7(1):190987.

5. Kauer T, **Joglekar S**, Redi M, Aiello LM, Quercia D. Mapping and Visualizing Deep-Learning Urban Beautification. IEEE computer graphics and applications. 2018 Sep 27;38(5):70-83..

1.5.2 Collaborative author contributions

List of papers where the contribution was significant, but were not led by the author. Contributions from these articles are not included in this dissertation

1. Bhatt S, **Joglekar S**, Bano S, Sastry N. Illuminating an ecosystem of partisan websites. In Companion Proceedings of the The Web Conference 2018 2018 Apr 23 (pp. 545-554). International World Wide Web Conferences Steering Committee.

2. Raman A, **Joglekar S**, Cristofaro ED, Sastry N, Tyson G. Challenges in the Decentralised Web: The Mastodon Case. In Proceedings of the Internet Measurement Conference 2019 Oct 21 (pp. 217-229). ACM.

3. De Simoni A, **Joglekar S**, Taylor SJ, Patel A, Duschinsky R, Coulson N, Griffiths C, Panzarasa P, Sastry N, Anand A, Evans MJ. Structure and dynamics of online patients' communities: the case of Asthma UK and BLF online fora..

4. Young AP, **Joglekar S**, Garimella K, Sastry N. Approximations to Truth in Online Comment Networks.

5. Boschi G, Young AP, **Joglekar S**, Cammarota C, Sastry N. Having the Last Word: Understanding How to Sample Discussions Online. arXiv preprint arXiv:1906.04148. 2019 Jun 10.

6. Magdy W, Elkhatib Y, Tyson G, **Joglekar S**, Sastry N. Fake it till you make it: Fishing for Catfishes. In Proceedings of the 2017 IEEE/ACM International Conference on

Advances in Social Networks Analysis and Mining 2017 2017 Jul 31 (pp. 497-504). ACM.

1.6 Thesis overview

We begin with my first case study on Online communities. In Chapter 2 (based on article 2), I examine how supportive communities evolve and sustain over a long period of time. I show presence of an anti-rich club effect on these support groups, which implies that experienced users are more interested in helping new comers rather than forming a clique of their own. I define a quantitative metric for "expertise" and show that as one becomes adept, one becomes more willing to help. All these insights point towards answers for **RQ1** and **RQ2**. In Chapter 3 (based on article 3), I look at structural signatures in the interactions between users in a supportive conversations. I show that mapping the conversational exchanges onto a topological structure, exhibits keen preference for local supportive motifs, which I call "anchored motifs". I discuss the utility of such a model of support conversation and draw parallels with the offline model of community support(Chapter 4) as per the mandate of **RQ3**.

In the second study, I investigate utility of perceptions of real world places through a crowd sourced rating of google street view images. As per **RQ4**, I develop models to extract the perception of the crowds using data driven inference methods(Chapter 4)(based on article 4). I then show that a general pattern of beauty in urban spaces can be learnt through a crowd sourced opinion and based on this finding, I develop a generative model to simulate beautification of urban spaces by using deep learning(Chapter 5)(based on article 4 and article 5). I validate the quantification of perception of real-world beauty using crowd validation. I contribute a way to use computer vision techniques to abstract out beautification process into explainable metrics used by architects and urban planners. The final contribution is a demo web application, that allows practitioners to examine and validate the utility of such a end to end system that captures citizen perceptions for urban design. These contributions are motivated by **RQ5** and **RQ6**. I close by enumerating the different research problems and future directions that my work would pursue as a early career scientist(Chapter 6)

Chapter 2

Background

If I have seen further it is by standing on the shoulders of Giants. - Isaac Newton

Studying signals in user interaction data, where the interactions are driven by affective triggers, has been an active topic of research [Pic03, PPNH07, CGHH12, AJS$^+$13]. As such it is worth discussing the different aspects in which the community of researchers have explored this area. The key aspects in which I would like to place my work is in terms of quantification of the subjective signals from data that originates from web scale applications. The two case studies in this dissertation attempt to quantify two distinct subjective properties, viz 1) Social support and 2) Aesthetic perception. Here we would try to first, explore the definitions of the concepts of social support and Aesthetic perception, and then examine the literature for methods, models and metrics.

2.1 Part 1: From communities

There has been a surge in the number of online communities, since the rapid adoption of social networks across the internet. The spectrum of types and applications of these communities is as abundant as the possible subjects discussed on the internet. These communities have become a dedicated space for online users to discuss about topical items. Often these communities take the shape of a forum, where topical threads are started by an original poster (OP) and a discussion commences on this post. The discussions could be in the form of a debate[1], a banter[2] or a peer to peer topical discussion[3]. In this dissertation, we aim to understand the nature of social support and the signatures of social support which can be quantified from these online spaces. To that end we aim to look at forums which have been self certified to be dedicated for hosting supportive discussions.

2.1.1 Online social support

According to Shumaker et.al [SB84] social support is defined as "an exchange of resources between two individuals perceived by the provider or the recipient to be intended to enhance the well-being of the recipient." A lot of work has been done in understanding how social support functions and influences people in distress in the offline world. For example, a meta review [DiM04] showed that adherence to a medical treatment is 1.74 times higher, if the patient hails from a cohesive family structure. Social support has also shown to be a crucial factor in the positive prognosis of patients suffering from chronic and long term conditions [SY06, PFR+00, BAH+86, CDSLS93, DS84, BCH+90]. All this literature evaluates social support from a psychological stand point, in that, it looks at how a patient/subject is perceiving support from its real world network (family, friends, doctors etc.) More over most of this work uses the qualitative frameworks and tools as a way to measure off-line social support.

Social support, or the perception of help received from others, is a widely studied as a psychological resource used to cope with distress. Social support is generally classified into one of the following categories viz. Informational, Tangible , Network or Emotional [CS92]. These categories measure the nature of social support along the idea of exchange of resources between the person in distress and the one providing support. For example, an informational support could involve pure exchange of valuable information about dealing with an issue. Whereas as network support could purely allow the recipient to acquire a wider network of people through the support giver or a support platform (think societies like Alcoholic Anonymous).

[1] www.kialo.com
[2] www.reddit.com
[3] https://healthunlocked.com/blf

The online world is very good at filling the gaps where offline support groups may fail. That is, the online groups, if designed in the right way, could provide essential informational, network, and at times emotional support at the click of a button. The internet makes both information and user networks, easily accessible. The utility of such spaces can be believed because of the evidence such as: online support groups being associated with quality of life [IKW09, Nam11, Cou05], improved control over additions [WW09], improved triaging with suicide ideation [CK17] and mental health issues [KGB+02].

Despite the interesting evidence, the common theme in most of these works is that they all take either a qualitative approach to understanding online support, or a language driven approach [DVZK+14, DSSBSM16, DSHF+17]. Most methodologies either utilize expert knowledge to dissect what is being said on these forums, or a natural language approach to understand the key language patterns on these forums. Either ways, the key missing bit in this picture is understanding how and where do the individual users fit in. How do they help in keeping the entire support community functioning. More over, we do not have any understanding about the structure of a supportive conversation online from the perspective of a user in distress. This is especially interesting, since there are obvious offline markers of a supportive conversation, whether it being a group, or a peer to peer setting. Not just that, but there are several therapy strategies like Communication accommodation [CG88] or group psychotherapy [Yal95]. So it is worth investigating these aspects of online support communities.

For this exact reason, **RQ1, RQ2 and RQ3** would make progress towards understanding the dynamics, local and global structures of online support communities.

2.1.2 Social networks, support, and metrics

Modelling and studying online social networks, through the lens of complex network theory and network physics has been an active area of research since the 1980s. The idea of looking at (offline) social structures as social networks was quite prominent in the fields of sociology [Sco88], but the most important leap in this field came with the advent of online social networks. This opened up new doors in measuring and understanding how humans form networks, and do so at scale through the medium of online networks [MMG+07]. The first of these works [MMG+07] explored the idea of looking at large scale user graphs to measure and evaluate a lot big picture attributes about online social networks. In that, they measure the long tailed nature of social links, the symmetric nature of social links and the overall sizes of connected components. A connected component in a social graph $G(V,E)$) – where V are the vertices or users and E are the edges between them – is a set of nodes $\{C\}$ where all nodes $\{c_i \in C\}$ have connected paths to all other nodes in $\{C\}$. The size of the largest connected component was often used as a proxy for the connectedness of a social graph [MSGL14, TMP12, WRP+94]. Other global scale metrics explored for understanding social networks were degree distributions [MPP+13, NWS02, KW06], clustering coefficient [Ops13, TOS+06] and centrality [OAS10, BMBL09]. All these metrics aim to look at how nodes(users) in a social graph group together or how do they interact with each other as the size of the network increases. These insights help network physicists model how information or a contagion diffuses in a connected community. However , despite the large number of interesting works that look at network structure, there have been limited progress in using these tools to understand the nature of social support in online networks.

In the offline world, there have been some studies that look at the ego network of a person of interest, to infer the nature of social ties they have. In that, they look at the transitive nature of social networks around a person [GCL09, HL71, LCB01]. This means, how many friends of a person, are friends among themselves. A completion of this triad – which implies that friends of my friends are my friends – is called a triadic closure. In the online world, the theory of social capital and triadic closures were operationalized in terms of triadic census [Fau07, Fau08]. The census simply profiles any given network by counting individual instances of the individual sub-graphs made of two, three, or in some cases four nodes. These sub-graphs are called motifs. The importance of triadic motifs in social network research has been stressed so much so that in the words of Holland and Leinhardt [HL77]

> “The essential issue of any notion of structure is how the components are combined, not the components themselves...this issue amounts to the proposition

that the lowest interesting level of structure...is the level of triples of nodes—the triadic level"

From the studies discussed here, the link between human to human interactions and their emergent network structure are quite evident. For this reason it is natural to extend this link to explore how perceived social support manifests in these interactions. In this dissertation, I aim to first quantify how network structures in support communities evolve, and second, discover the signature structural properties of these interactions over support communities.

2.2 Part 2: From crowds

The second part of the dissertation explores new methods to use crowd's opinions, in order to build models of our subjective perception of urban spaces. The idea of using crowd based annotations or opinion mining has been popular and has been exploited in the recommendation systems literature for a while. Here I would introduce some background about the idea of crowd sourcing, how it applies to quantifying the subjective, and how it has been used to understand urban spaces and cities.

2.2.1 Crowd sourcing and the subjective

Crowdsourcing has been an important part of the computer science research for the past decade. The idea of crowdsourcing was first brought to attention by Jeff Howe [How06], where he introduced the idea of a logical equivalent of outsourcing – which is sending the jobs outside an area, where labour prices are competitive– but for more transactional and atomic tasks. This idea was quickly adopted by the academic community, right after the inception of services like Amazon Mechanical Turk or Crowdflower [PCI10].

The natural extension of this new method was to use it for annotating large quantities of data. These annotations generally dealt with labels of objective nature, such as objects [VPR13], relationships between objects [KZG+17], or annotating textual data like named entities [FMK+10].

The key idea behind crowdsourcing is to get a cognitive input about unlabelled data by incorporating opinions of hundreds of "crowd-workers" exchange for money. This is done in order to annotate the data with the most accurate objective labels, that come from a human annotator. This data then becomes the training set for a downstream machine learning algorithm . These algorithms would then learn the task of classifying unseen data into their respective correct labels.

But the very fact that a cognitive process is at the root of the annotations implies that this framework can even be applied to subjective properties, provided that the annotators can arrive at a consensus, and we have large enough samples. The most appropriate use case is that of annotating expressions or affects of humans in videos or images. Despite the subjective nature of perceived affect, most neuro-typical humans tend to agree on what constitutes the expression of anger, sadness, happiness , disgust etc. In that spirit, several studies [TMM16, KAH$^+$16, KV16] tried to build machine learning models that could detect emotions from facial expressions, using data annotated by the crowd.

Apart from building a model of human affects from faces, crowdsourcing was also applied to the area of quantifying the actual affective stimuli in content. That is it tries to quantify the intangible property of a content that stimulates evocation of a particular emotion in the consumer of that content. For example, it is worth asking “Which emotion does an image of a sunset evoke in the human seeing it?”. This question goes one level deeper by trying to quantify that which evokes positive or negative sentiments. To that end Sentibank [B$^+$13] explored this idea by training a deep learning model on Flicker images which were annotated for evoking positive or negative sentiments. The same team extended it to analyse how the evoked emotion changes as a function of culture and language [PRT$^+$16]. Indeed they found that these evoked emotions are also dependent on the language, culture, and other social properties of the annotator. Although these works exposed some limitations in the approach of quantifying the subjective, they also showed that by and large, these techniques work if the data is large enough and there is considerable consensus among the annotators on the topic of the annotated subjective property of the data.

2.2.2 Crowds and the cities

So far, the most detailed studies for quantifying the perceptions of urban environments and their visual appearance have relied on personal interviews and the observation of city streets: for example, some researchers relied on annotations of video recordings by experts [SR04], while others have used participant ratings of simulated (rather than existing) street scenes [LH12]. But since the advent of services like the Google street view and Open Street Map, the Web has now been used to survey a large number of individuals. Place Pulse is a website that asks a series of binary perception questions (such as 'Which place looks safer [between the two]?') across a large number of geo-tagged images [SSH13]. In a similar way, Quercia *et al.* collected pairwise judgments about the extent to which urban scenes are considered quiet, beautiful and happy [QOC14] to then recommend pleasant paths in the city [QSA14]. Another study [SPM15] presented the annotators with a 10 point scale, which they would use to score a place(Street view) for its aesthetic beauty. All these studies relied on the crowds, in that the annotators were completely disconnected from each other, and their ratings were purely based on their exposure to the image or an urban scene. An important caveat here, as in case of multi lingual sentibank [PRT+16], is that the cultural and social background of the annotator would play a role in how they perceive an urban scene. But on average, these annotations proved very useful in understanding something as subjective as the perception of safety, beauty, or memorability in urban spaces.

This can be indicated by the fact that lately deep learning techniques have been used to accurately predict urban beauty [DNP+16, SPM17], urban change [NKR+17], and even crime [DNVZ+16a, AERA14]. Recent works have also showed the utility of deep learning techniques in predicting house prices from urban frontages [LSSGR18], and from a combination of satellite data and street view images [LPR19].

All the studies discussed above were successful in quantifying the subjective properties of an urban scene using predictive machine learning models. But there is a significant gap between predicting and explaining the prediction in order to guide interventions. This explainability problem is prevalent in almost all applied machine learning system. In this dissertation, I attempt to make some progress on the front of explaining the reasons behind perceiving an urban scene beautiful or ugly. These explanations also come in the form of urban design metrics, which can guide interventions from the practitioners of this field.

2.3 Discussion

The title of my dissertation enumerates the two regimes –communities and crowds– under which I explore the problem of capturing the signatures of perceived subjective properties, using customized metrics and models. The over arching thesis has always been understanding how human perceptions guide our actions on the web scale, and how these actions leave behind traces of the subjective triggers.

In the following chapters, I will discuss the different methods, metrics, and models that are developed in order to make progress in answering the central thesis of this dissertation.

CHAPTER 3

FROM COMMUNITIES: THE ACTORS OF SOCIAL SUPPORT

> *"The original idea of the web was that it should be a collaborative space where you can communicate through sharing information... In an extreme view, the world can be seen as only connections, nothing else.“* - Tim Berners Lee[BLF01]

Attention budgets pretty much govern how we as consumers interact with online social networks. It has been shown that the dearth of this budget, promotes an engagement behaviour that prioritizes perceptive features and immediacy in the content [JSR17]. The scrolling user interface of platforms like Instagram and Facebook, allow mere seconds to decide whether a particular content is worth the user's attention [Eik17].

However, there is a whole breed of online social networks, which aim at bringing the offline sense of networking, online. These networks are mostly designed around a specific purpose like technical discussions[1], subject specific questions[2] or simply around hobbies like knitting[3] or art[4]. These communities embody the true essence [BLF01] of the internet, in that they strive at making geographical distance secondary, to the act of social networking and information sharing.

Definition 3 *According to the seminal work by Shumaker and Brownell [SB84], social support is defined as "an exchange of resources between two individuals perceived by the provider or the recipient to be intended to enhance the well-being of the recipient."*

[1]www.stackoverflow.com
[2]www.stackexchange.com
[3]www.ravelry.com
[4]www.artween.com/

Under this construct , these communities are apt Petri dishes to study the signatures of the online social support. Once you could quantify the social support signatures in terms of computable metrics, platforms could then empower the participants of these communities and design interventions to curb negative behaviour like trolling.

In the context of this dissertation, I wanted to know how signatures of a perceived entity like social support, manifests on these formal social networks. More specifically I develop methods and frameworks, to extract the signatures of perceived *social support* on communities designed around users who have underwent or are undergoing physical or mental distress. These methods and frameworks could bear significant potential impact on the health and utility of these forums. The first step us to understand the structure and utility of these communities. More over, having a primer on these communities would help the reader get an idea about the methods developed over the course of this dissertation.

3.1 Primer on online health communities

Recent work has proposed that online communities have the potential to influence health and health care sectors. Recent studies have suggested that the participation of people with long-term conditions (LTCs) in online communities (1) improves illness self-management [AVKR16], (2) produces positive health-related outcomes[5] [MC12, PS15] , (3) facilitates shared decision-making with health care professionals [BC11, IABS+17], and (4) may even reduce mortality [HBCF16].

There is also evidence that self-management support interventions can reduce health service utilization [PRS+14, TPE+14]. This is especially a crucial point as the world health services are facing the brunt of an ageing population.

Online communities have experienced an upsurge in popularity among people with chronic respiratory conditions such as cystic fibrosis [KM16], asthma [SLM+11], pulmonary hypertension [MMA+13] and chronic obstructive pulmonary disease (COPD) [WB13]. More than 15 million people in England suffer from a long-term condition or disability, and they account for at least 50 percent of all general practitioner appointments[6]. Thus, assessing how

[5]https://bit.ly/2FLcs1F
[6]https://bit.ly/2EVFs9v

these online communities function, evolve and provide perceived support, can have important implications for health care sector. More so, understanding the dynamics of these online communities, have actual repercussions on how the platforms that host them, could become a better resource of self-management of LTCs.

On average, one in four people with an LTC who use the Internet tries to engage online with others with similar health-related concerns [Fox11]. In particular, it has been suggested that the value of participating in an online community lies in the possibility of gaining access to a range of people and resources quickly, easily [AH00], and anonymously [PS15], as well as obtaining tailored information and emotional support [AFGG15, DSHF$^+$17, SC16, Cou05, DSSBSM16]. However, most of this evidence comes from qualitative studies, whereas only recent years have witnessed an increasing interest in quantitative assessments of online communities as intervention mechanisms.

The potential future integration of online health support systems with formal health care provision should be underpinned by a better understanding of how they are used and by evidence of their effectiveness. Indeed, as suggested by the Medical Research Council [CDM$^+$08], integrating online support systems with the more traditional health care provision would require the identification and comparative assessment of potential alternative intervention mechanisms.

In this chapter, I aim to uncover and understand how these communities function,evolve, and operate in the role of a supportive entity, from a macroscopic perspective. Further it is worth knowing if there exists any particular mechanism, by which individual users evolve into the role of support givers.

Ultimately, once we understand the mechanisms of peer support, we can proceeded in looking for discriminative signatures of these supportive exchanges. In this chapter, we would illuminate the answers for **RQ1** and **RQ2** using a large scale dataset of a pulmonary illness support group. The questions are:

RQ1 *What dynamics of support communities help them thrive?*

RQ2 *What differentiates* ***users*** *on support communities from generic ones?*

3.2 Dataset and properties

The data was collected from HealthUnlocked[7], the online platform provider of the Asthma UK and British Lung Foundation communities. Registered users can choose to either write posts publicly or send private posts to one another. In the latter case, posts are shared between 2 users only, whereas when posts are written publicly, a large number of users can become connected through threads of posts. A thread is a series of posts made on one root post, as a response to the root, or as a response to one of the responses to the root. This tree-like structure of posts can evolve indefinitely between posters. Only posts that were shared publicly were collected and analyzed. For this study, user identifiers (IDs) were anonymized by the HealthUnlocked platform, and no demographic information was collected. The data set included posts and their metadata (ie, the anonymized user ID numbers), user roles (eg, user, administrator, or moderator), date of posting, the hierarchical level of the post within the corresponding thread, and the dates in which the users joined and left the community. Both communities were moderated, and HealthUnlocked moderators (identified through metadata linked to posts) were included in the analysis to assess their contribution and compare it with other users. Online communities on the HealthUnlocked platform benefit from additional functionalities compared to other online forums, such as built-in patient groups that moderate the content. In particular, the content accessed by users is tailored to their interests, and profiles highlight users' condition, chosen community, medications and treatments they use or find interesting. No data were collected on participants' characteristics, though only people declaring themselves to be older than 16 years were permitted to create an account and take part in the online communities. Table 3.1 summarizes the salient features of the dataset used for this work.

[7]http://www.webcitation.org/70Y10rppl

Dataset Properties		
Property	**AsthmaUK**	**British Lung Foundation**
Time span of data	02/03/2006-06/09/2016	13/04/2012-06/09/2016
Total Time (weeks)	548	230
Total number of posts	32,780	875,151
Percentage of posts with at-least 1 reply	87.3%	93.1 %
Total number of users	3345	19,837
Users who contributed > 1 posts (%n)	1053 (31.5)	7814 (39.4)
Users who contributed exactly 1 post(%n)	331 (31.4) 722	1186 (15.2)
Registered users who never posted (ie, lurkers), n (%)	2292 (68.5)	12,023 (60.6)
Number of posts per user, $\mu(\sigma)$	14.2 (55.0)	66.9 (75.1)
Number of posts per users who posted >1, median (min - max)	5.1 (2-1068)	8.0 (2-8947)
Number of posts per users who posted >1, mean (SD)	20.4 (65.6)	88.1 (458.6)
Posts contributed by top 1% users by activity, n (%)	10,457 (31.9)	426,198 (48.7)

Table 3.1 The summary of salient attributes of the data used for this work

The data sets span, respectively, 10 years for the Asthma UK and 4 years for the BLF communities (see Table 1).

Despite the shorter time span, as a result of the larger number of users, the number of posts in the BLF community was higher than in Asthma UK, namely 875,151 compared to 32,780 respectively. Moreover, BLF users wrote a higher number of posts per user and were connected with a higher number of other users when compared with people in the Asthma UK forum (see Figure 2). In both communities, 60%-70% of registered users wrote no posts (ie, they were lurkers). Users who wrote more than one post contributed with a median of 8 (range 2-8947) and 5 (range 2-1068) posts in the BLF and Asthma UK communities, respectively.

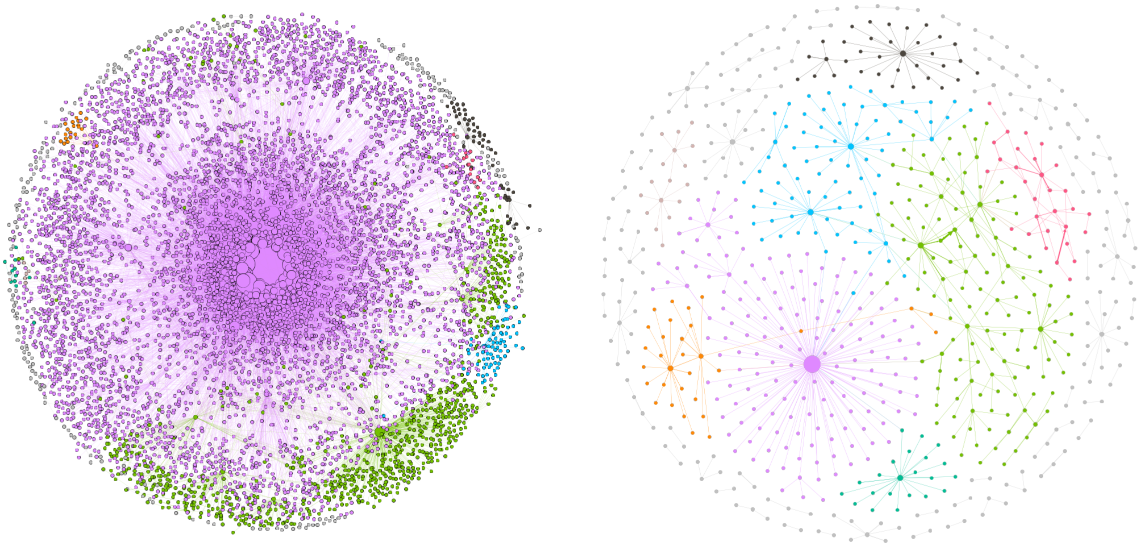
The number of official moderators among the highly active users was negligible; there were no moderators in the top 5% contributors to BLF and only 2 in the top 5% for Asthma UK. Thus, our network analysis predominantly reflects content originated from registered users. This also means that moderators on these forums have more of an observatory role and do not engage in active support.

When classified according to posting activity (ie, number of posts written to the forum), the top 5% users contributed to a substantial proportion of all posts: 58% and 79% in the Asthma UK and BLF communities, respectively. In the context of this thesis, *Superusers* were those who made high number of connections with other users across the lifetime of the community.

3.3 Interaction Graphs

To understand the reason behind how these communities thrive and in order to quantify the conversation structures, I convert all the message exchanges into graphs, where users are represented by nodes and messages are represented by edges between users. More formally imagine a directed graph $G(V,E)$ involving a set of users $V_i \forall i \in N$ where N is the total number of users interacting on a health community. For every message exchanged between a user i and a user j we create an edge E_{ij}. The complete community would form a global graph based off total interactions between all pairs of users which we call a global graph G_g. Similarly we may decide to only consider the users and messages exchanged across one particular thread discussing a particular issue. Such a graph is called a thread graph G_t.

These graphs are the abstractions of how users interacted on the community either around a particular query (Thread graphs) or over all as a part of the bigger community (Gobal graph). To understand the behaviour of these users, I evaluate several metrics on these graphs to understand the utility of these communities in terms of activity of sharing and support. This abstraction also makes it feasible for us to investigate how a particular community grows with time and how particular users evolve with time.



(a) (b)

Fig. 3.1 Global graphs prepared from Asthama UK community3.1b and BLF community3.1a. The size of the node corresponds to the degree of the node and the color corresponds to the community membership

3.4 How do support communities thrive ?

This question sets the primer about the peculiarity of support communities. Do these communities have enough interest and activity, to sustain over long periods of time? Are all users equally involved in the vitality of the community or is this a group of users? How important are these users?

3.4.1 Activity Metrics

To calculate the activity patterns of users on these forum, we first work with the most basic of proxies, which is the weekly/daily activity. We arrive at it by calculating the amount of messages exchanged in a community across the whole life cycle of the data. This metric would expose how much activity is happening on a daily or weekly basis on a particular community. It is worth noting that this activity pattern, would also shed light on how users are engaging with the community. A continuous engagement is good for the vitality of a

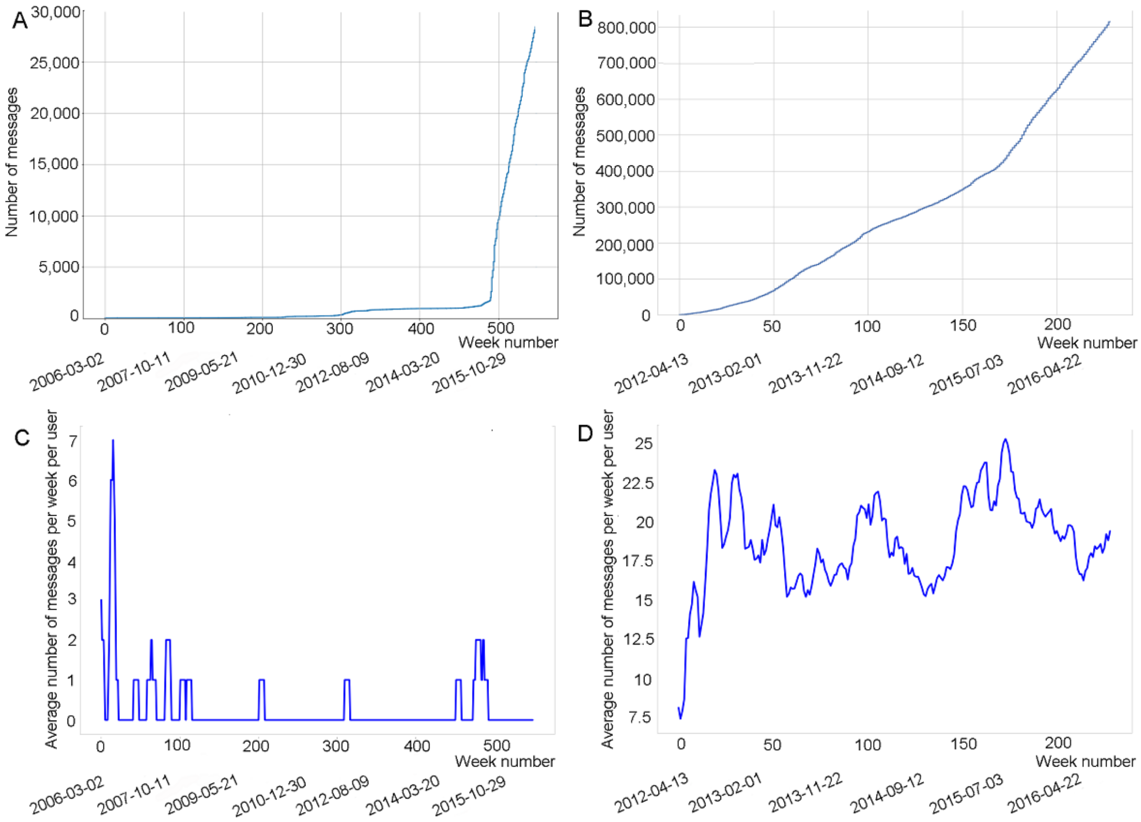


Fig. 3.2 Cumulative distributions of the number of posts as a function of time (weeks) within the Asthma UK (A) and the British Lung Foundation (B) communities. Calendars dates are reported below as week numbers (since the inception of the community). Panels C and D illustrate the average number of posts per user per week within Asthma UK and British Lung Foundation, respectively

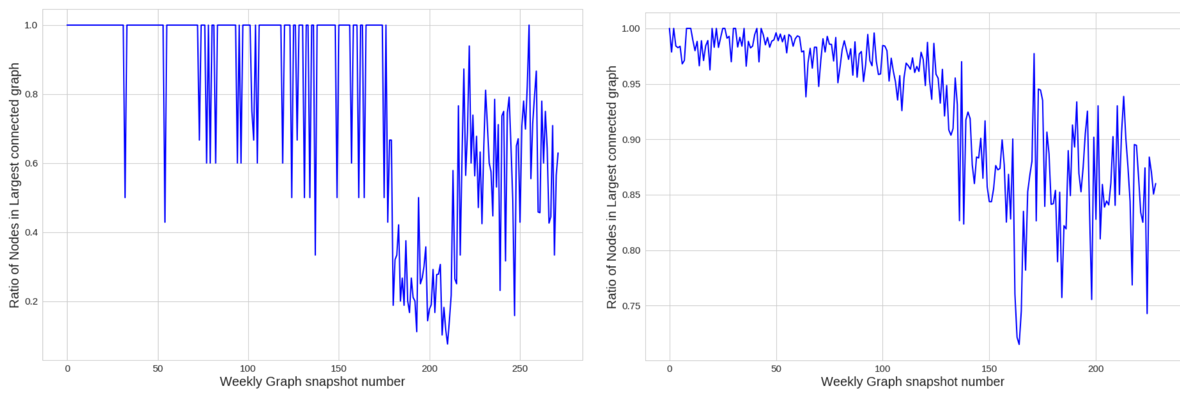
community, however if a community revolves around purely functional interactions, then you may see a bursty nature of communication [PB15]. From these basic analysis, it was quite evident that the BLF community was more active of the two, in that, the community exhibits a consistent engagement of the users across the lifetime of the data. Asthma forum however shows a bursty activity, despite being more than twice as old as the BLF community. The trends can be see in Figure 3.2.

Remark 1 *It is worth noting that as the activity on the community increases, you do see an increase in fragmented reply networks, which means concurrent discussions are happening with disjoint set of users interacting.*

3.4.2 Community resilience

The activity patterns does not tell the real story about the cohesiveness of any particular community. To answer the **RQ1**, it is first worth asking how the user interactions bind the community together. We would like to know if the user activity is highly concentrated or is it covering a large fraction of the user base. More so, it is worth asking if there are any special users who bear the mantle of providing support. This can be observed from the topological properties of the interaction graph. From table3.1, it is evident that a minority of users are generating a bulk of data on these communities. E.g. the top 1% users by activity contributed 32% posts to AsthamaUK community. Such level of activity makes these users extremely important in understanding the dynamics of support on these communities.

Cohesive conversations



(a) (b)

Fig. 3.3 Fraction of users that are part of the largest component as a function of time (weeks) for Asthma UK 3.3a and the British Lung Foundation 3.3b.

To understand the first aspect of the community's resiliance, I examine how is the coverage of communications between the users, on a weekly basis, given that messages are exchanged between the most active users within that week. To do so, imagine a sorted list of message interactions over a particular time period T_k, sorted in chronological order defined as $L_k = [E_{ij} \forall i, j \in N]$, where E_{ij} is a message between user i and user j, with N total users being active in a given time period T_k. Now imagine this time period T_k is of 7 days . I calculate such K lists for the K weeks the community has been active. For each such list, I induce a graph $G_k(V,E)$ such that the nodes in V are the active users in that particular list, and the edges in E are corresponding to the messages exchanged in the list L_k between any two users.

Definition 4 ***Largest Connected component*** *: A largest connected component of a Graph $G(V,E)$ is the largest possible subgraph $G_L(V_L,E_L)$ of G, such that each node in G_L has at least one valid connected path to every other node in G_L*

Now for each such graph G_k I calculate the largest connected subgraph $G_{\theta_k}(N_k,E_k)$ such that all nodes in N_k have at least one path between them. Calculating the fraction $\frac{N_k}{N}$ would give us the total fraction of users who are part of the same conversation network for a given week. After calculating and plotting these fractions across a total of 250 weeks for each community, we see that whenever there is an activity on these networks, almost always, the active nodes belong to the largest connected sub graph. This implies that activity on support forums is cohesive and even if bursty at times, is all encompassing with the users.

Fragile communication structure

Despite the exchange on a weekly basis is quite cohesive, it is pertinent to understand the resilience in terms of user responsibility in helping, in order to examine the health of such a community. Moreover, I want to know if the conversation network is held together by a more or less uniform contribution of nodes, or if there is a skew in the responsibility of nodes. This can be tested by using the sensitivity analysis methods, popular in the network science [BDSZ16, AJB00], which measures the network's capacity to diffuse information as you remove nodes based on certain property. In our case, we want to understand the importance of the *Superusers*, or the users who are disproportionately more active. Hence

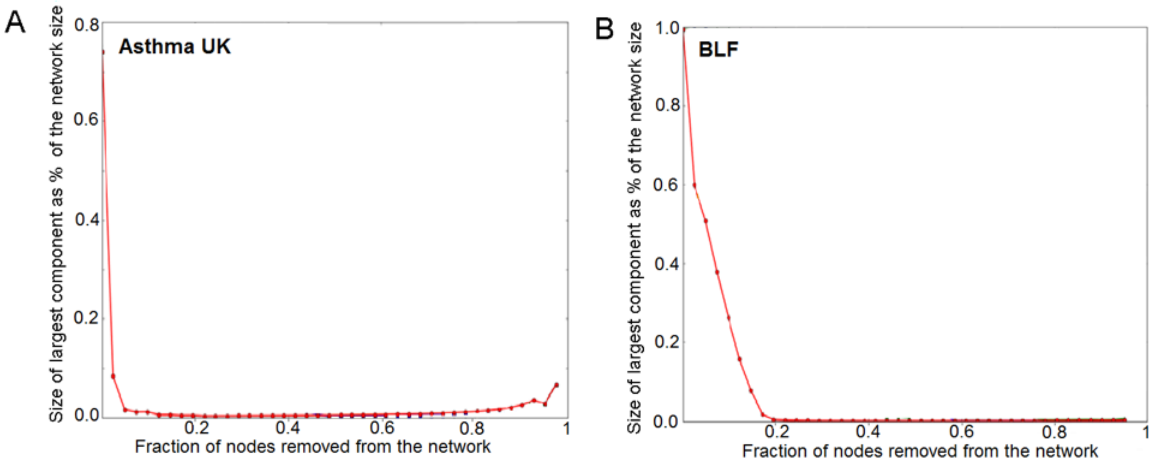
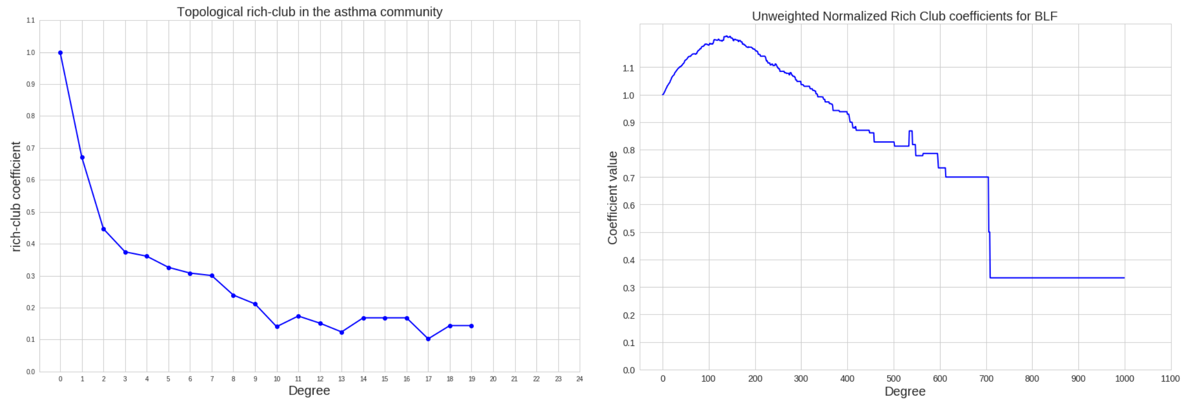


Fig. 3.4 Results of progressive removal of nodes based on connectivity. Both communities collapse drastically, in terms of connectivity, with BLF showing a little more resilience

we begin by first sorting all the nodes in the macroscopic graph $G(V,E)$ in order of their degrees. The degree of a node in the global graph is proportional to the diverse set of users that node has communicated with, over the period if the community's lifetime. We then start removing nodes from the top, by progressively removing nodes in increments of 1%. I then compute the size of the largest connected component G_k and compute the ration of number of nodes in G_k as compared to the original global undisturbed network. Figure 3.4 shows the performance of global graphs of both the communities to this attack. It is worth noting, that what we observe is that a top 10% nodes by activity are responsible for most of the cohesive connectivity of the community. This also means that the top 10% of these nodes have the most diverse connections in terms of number of users contacted. This gives hope to health care industry, since these nodes can act like efficient information diffusers, if used in a targeted fashion.

Anti-rich conversations

The "rich-club" coefficient is a metric designed to measure the extent to which well-connected users tend to connect with one another to a higher degree than expected by chance [CFSV06]. To this end, for each value k of a node's degree (ie, the number of other users a given user is connected with), we computed the ratio between the number of actual connections between



(a) (b)

Fig. 3.5 Plots of rich-club coefficients for each viable degree in the respective communities.

nodes with degree k or larger and the total possible number of such connections [OCPR08]. We then divided this ratio by the one obtained on a corresponding random network with the same number of nodes and degree distribution (ie, the probability distribution of the degrees over the whole network) as the real network, but in which links were randomly reshuffled between nodes.

Formally let $G(V,E)$ be a global graph representation of the community. Let $V_{>k}$ be the set of vertices in the graph having degree higher than k. Let there be $N_{>k}$ such vertices having $E_{>k}$ edges between them. In such case, the rich club coefficient for degree k in the graph G is given by

$$\phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k}-1)} \tag{3.1}$$

In this equation $\frac{N_{>k}(N_{>k}-1)}{2}$ represents the maximum number of edges possible between $N_{>k}$ nodes. These coefficients are highly dependent on the size of the network, which makes them hard to compare. So I normalize the network by comparing against a random null model of rich-club coefficients $\phi_{rand}(k)$. This is obtained by generating an ensemble of random networks, each having the same degree distribution as that of G, but with links randomly placed. The ratio $\frac{\phi(k)}{\phi_{rand}(k)}$, gives us an un-correlated trend about the rich-club effect in G.

Thus, the rich-club coefficients may take values lower or higher than 1, depending on whether the real network has a higher or lower tendency to coalesce into rich clubs than randomly expected. In particular, networks that display a high rich-club coefficient (ie, greater than 1, are also said to show a "rich-club effect," namely the tendency to organise into a hierarchical structure in which highly connected nodes preferentially create tightly knit groups with one another [MdFCC07]. .

Remark 2 *It is worth noting that most previous studies in base lining rich-club effect in technical and real world networks have shown a propensity to create rich-clubs. Thus these networks exhibit exclusive clubs of (topologically) rich nodes, as illustrated in previous work [ZM04, CFSV06]*

What we observe from this analysis is that communities in general have a less than 1 rich-club co-efficient value for a spectrum of degrees k. This means, rich nodes are exhibiting an anti-rich behaviour, where nodes which have a higher degree, prefer engaging with new nodes with lower degree. This implies an active information exchange from a well connected node to a sparsely connected node, which follows according to the definition of social support(definition 3).

3.5 How do we quantify support on these communities?

Once I establish that these support communities are thriving and are providing what seems to be an active supportive environment for the patrons, it is worth delving into the analytical methods for quantifying these supportive interactions. More so we would like to have concrete metrics that characterize a given community as a supportive one. To do so we need to understand how are the users on these communities driven to help each other, and whether there is a correlation between the "richness" of a user, as defined in previous section, and its propensity to help. More so we would like to know how consistent are these so called "rich" users in providing support.

3.5.1 Propensity to help

We would like to understand how users on support communities, as a group behave as they become more seasoned. Fortunately, there is an approximate way for us to capture a user's role as a support seeker and as a support giver. As described in Section 3.2, the forum activity consists of a root poster, asking a question to the forum board, and the members responding to that question in a cascaded fashion. These responses, along with the original question constitute what is called as a *thread*. To that end, we define the following two roles on these communities[8]

Definition 5 *Support seeker: a user who begins a thread by posting on the forum, a question, or a query, to which others may respond to.*

Definition 6 *Support giver: a user who responds to any post by a support seeker.*

Using these definitions I aim at modelling the statistical propensity of someone being a support giver or a seeker, as a function of their "richness". We first begin by calculating across the dataset, the average number of questions per user and answers per user, by finding the mean number of questions or answers posted by any user on the forum. We consider an expected probability of answering a question by a user as P_a as 2/3 and the probability of posting a question as P_q as 1/3. With this information we modify the definition of "Z-score" to quantify the expertise, used by Adamic. et. al [ZAA07] to arrive at the expression of expertise in the context out our support community.

Proof 1 *Consider a Bernoulli process for a user to choose to answer or post a question on the forum, with asymmetric probabilities for answering (P_a) and posting a question (P_q). For any user i the total number of posts n_i are the sum of total number of questions posted q_i and answers posted a_i and $n_i = a_i + q_i$ For a Bernoulli process the variance for the whole forum is given as:*

$$\sigma_{forum} = \sqrt{nP_a(1-P_a)}$$

[8]There are other ways to qualify someone as support giver/seeker, mainly using language sturcture, but here we consider only the bare minimum requirement to be considered as one, using the position in conversation structure

$$\sigma_{forum} = \frac{\sqrt{2n}}{3}$$

Similarly the mean for this process can be written as :

$$\mu_{forum} = nP_a = \frac{2n}{3}$$

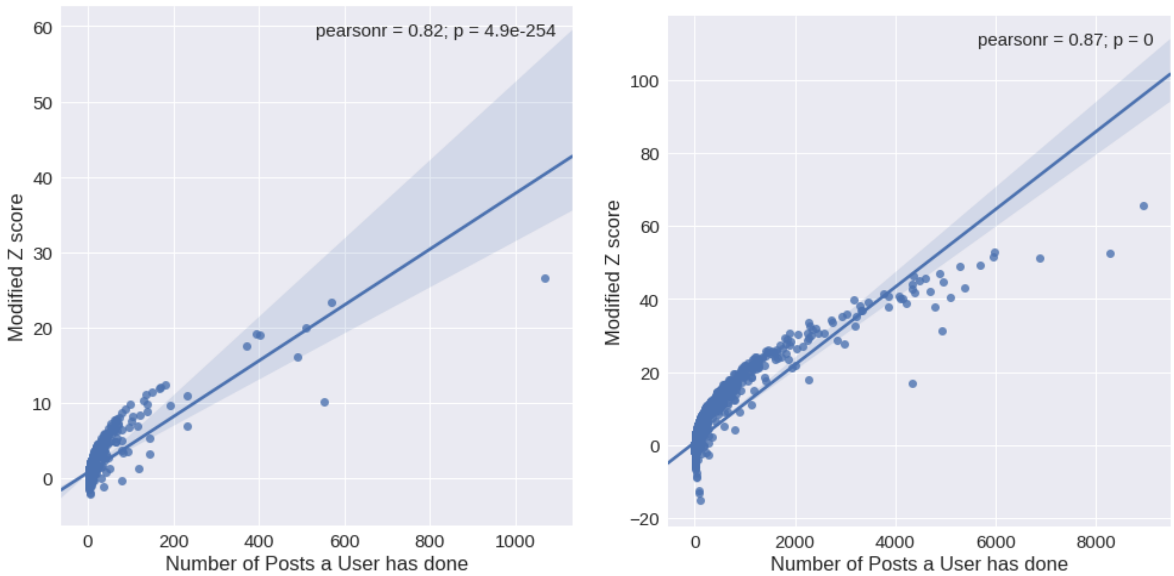
Z_{score} *of a random variable* X *is defined as*

$$Z_{score} = \frac{X - \mu}{\sigma}$$

Substituting the values for σ_{forum} *and* μ_{forum} *inside the expression for* Z_{score} *we arrive at the modified Z-score as*

$$Z_{score} = \frac{a - 2q}{\sqrt{2(a+q)}} \tag{3.2}$$

Equation 3.2 depicts the modified notion of Z-score for the question answering process of our support community. I calculate this particular metric for each user in both the communities based on their posting history.



(a) (b)

Fig. 3.6

I then find the correlation between a users Z-score and the total number of posts a user has done in their lifetime on the forum. Figure 3.6a and Figure 3.6b shows the results of this analysis for both the communities. It is quite evident, that as the users become more seasoned and post more actively, they are more likely to answer on questions rather than post new ones. This also implies that based on the rich club results from Section 3.4.2, these communities are thriving not only for the "rich" users, but also for the sparse users. Users on these communities are more open to new members and provide active support to them. Developing metrics like this makes quantifying whether a particular community works for the subscribers a tractable problem.

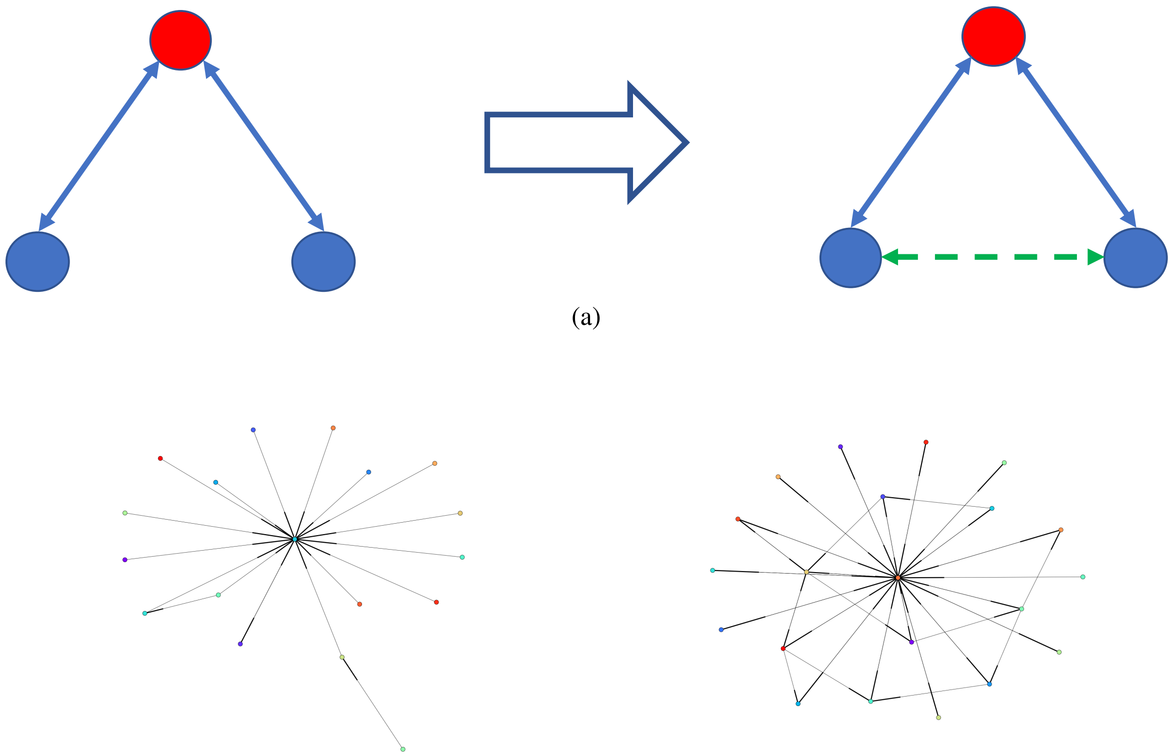
3.6 Rich users and structural holes

One of the key aspects of utility of any social network is driven from the social capital offered as a result of the subscription.

Definition 7 ***Social capital*** *(Sociology): Social capital is defined as those features of social structures , such as interpersonal trust and norms of reciprocity and mutual aid, which act as resources for individuals and facilitate collective action [CDSLS93, Col88]*

It is common to quantify social capital in the context of social networks, by looking at structural holes, or unmet potential social links in the network. This is where ties between otherwise unconnected neighbours are filled in, sometimes called as closures, thereby benefiting the broker and the two neighbours by adding an extra link for information to diffuse. Such mechanisms have been studied in the sociology literature for decades. Work by Granovetter [Gra77] explored these structural holes and proposed that they are detrimental for efficient diffusion of information and resources in social networks. He also at times called these the "forbidden triad", referring to their propensity to close up. Such closures are, according to Ronald Burt [Bur04, Bur09], necessary for information brokerage, and at times directly equate to social capital of these broker nodes. In our case, as so much evidence has shown that the brokers of social support are often the "rich" nodes or the super-users, we would certainly want to investigate how these agents affect the local cohesion and structural holes.

Till now we looked at the global macro structural properties of this support graph using the global graphs G_g, where we look at the user's interactions with other users across the lifetime of the community. But often the supportive interactions happen in a lifetime of a single thread, revolving around a topic or query. So to examine the effect of the rich users on the social cohesion, we correlate the total number of posts done by super users on any given thread, to the amount of closed triangles found in the corresponding thread graph G_t.



(a)

(b) (c)

Fig. 3.7 Figure 3.7a shows an example of closure among three nodes, where a structural hole between a cluster of three nodes is closed by addition of the green link. Figure 3.7b shows a thread level interaction graph showing lots of structural holes between participating nodes. On the other hand Figure 3.7c shows an example of a thread level interaction graph where a 'rich' user has contributed multiple times. This graph also shows more closures

The resultant scatter plot can be seen in Figure 3.8, where we can see a net positive correlation of 0.44, with a very low p-value. This means there is a general trend of higher triadic closures in a conversation, with the amount of rich user participation.

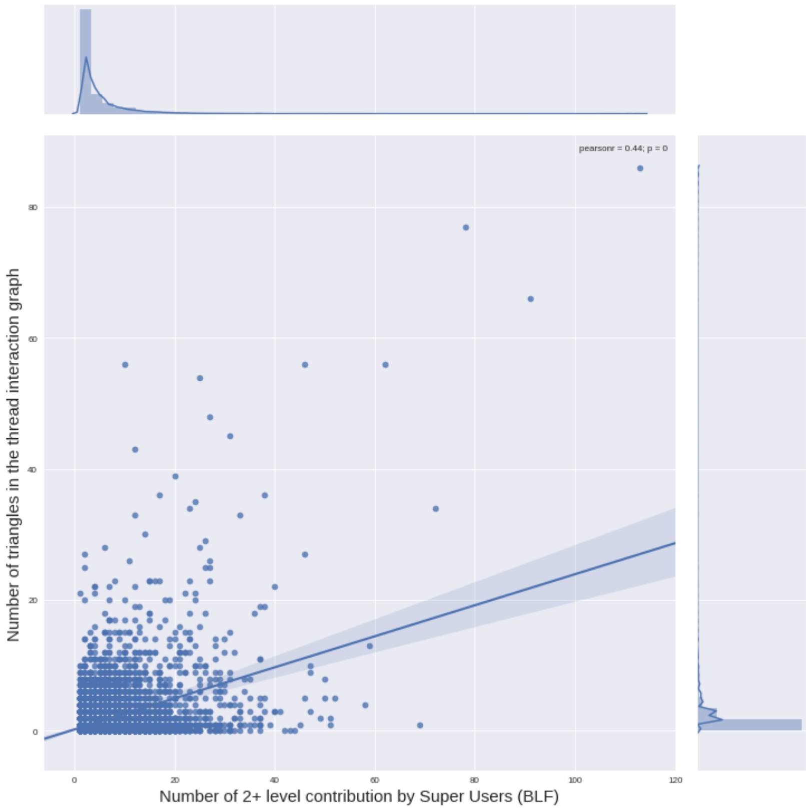


Fig. 3.8 Scatter plot between the number of closed triangles in a conversation thread, with the total number of posts done by rich users in that thread. A positive correlation of 0.44, with a p-value of 0 is observed

The research questions asked at the beginning of this chapter, required us to look at these communities from the perspective of its users. It was evident, that these communities do provide a supportive space. When looking at macroscopic structure of a support network, we find that certain groups of users are highly crucial for the vitality and survival of the whole exchange process. But we also observe that these groups are in flux, and users tend to become more supportive with experience. The idea of perceived support stems from the fact that the user in distress is not only getting the instrumental information, but also benefits from the social capital of the allied users. We observe that supportive users tend to have a positive effect on the social capital of a conversation, promoting more social cohesion.

Social support is a multi-dimensional perceived entity, and I am attempting to capture the mechanics of support through online conversations. It has been an active area of study from a sociological and psychological perspective. There is immense value in understanding how social support thrives on the internet, especially because of its potential of relieving some burden from the ever so burdened health services. The work till now did focus on the global interactions of support givers and seekers on a community. However, the actual action of peer support happens over a single thread of conversation. Aggregating these conversations over a period of time, allows us to look at how users behave by and large across different conversations. What is interesting and important to understand is how can we measure the act of peer support in an individual supportive conversation. How does a supportive conversation evolve, and are there any peculiar signatures we could observe in comparison with generic conversations? For this reason, it is important to look at a finer granularity of interaction, what we term as *mesoscopic* view of the conversation. In the following chapter, we would do exactly that, and unravel the discriminative patterns in meso and macro representations of conversation threads.

CHAPTER 4

FROM COMMUNITIES: SIGNATURES OF SOCIAL SUPPORT

The global structure of social support communities, allude to the fact that these communities are exhibiting some mechanisms different than generic communities. They also seem to have an effect on how people perceive support when in distress. In the past chapter, we saw the behavioural signatures of support groups. The utility of 'rich' users who are the most active and promiscuous when it comes to responding. The dynamic of evolution of users from support seekers to support givers. We also saw the way brokerage works in providing closures in the conversation threads, adding more closed triangles between people as 'rich' or expert users contribute. These results make it relevant for us to investigate the signatures of social support when it comes to individual conversations on the web. Quantifying how support manifests in terms of structure in conversation, might gives us more insights in how online support can be utilised properly for health care. This may also point to the features to look at, if we need to automate the process of quashing online toxic behaviour. Inspired from the DIKW pyramid, we have now experienced the transformation of raw social network interaction data, into abstractions that bring out peer to peer interaction patters as well as the structure of dialogue between users. We have also extracted the knowledge about emergent behaviours of users using topological metrics. In this chapter we will look into the the macro and the micro signatures of a supportive dialogue on these communities.

A valid criticism of the results from Chapter 3 is that the results were driven by a support community which is highly moderated. This sort of moderation is often absent on more popular open platforms. One such popular and open platform is Reddit[1]. Reddit offers its users, anonymity and very limited moderation. The limited amount of moderation comes from volunteers appointed by a particular community, or in the case of Reddit, a sub-reddit. A sub-reddit is a sub-section of Reddit, which is a topical community, where only topical conversation threads can be posted. Any user can subscribe and post to such a topical sub-reddit, provided they abide by the simple rules of the sub-reddit. For example, if a sub-reddit deals with politics, like r/politics, one must make sure that the threads posted on this subreddit deals with political subject matters. There are several sub-reddits on this platform that deal with peer to peer support. A few examples include r/SuicideWatch[2] , r/Depression[3],r/Bipolar[4]. Utility of these communities in providing a support, and a place to vent have been explored before [DD14, DCKD$^+$16]. However, in the context of my dissertation, I want to understand how such peer to peer support mechanisms are manifested in the dialogue structure. To do so, I need to look at these conversation threads from a over-all structural perspective(macro) as well as from a user-centric local perspective(meso).

4.1 An argument for studying mental health forums

The new platforms like facebook and reddit have become so ubiquitous, that some research suggests that they might be affecting our broader psychological state [CLPS12]. But on the positive side, studies have also proposed different ways in which this medium could be used for measuring and intervening in the matters of mental health [DCKD$^+$16, DD14]. These platform define the way humans interact in the present age, and developing an understanding of their effects and utility is of high importance. Reddit is one of the largest online communities which contains a number of topical sub-communities. On this platform, several subreddits are specifically tailored to mental health-related topics, such as *depression*, *anxiety* or *alcoholism*. These fora offer a unique opportunity to study the way people describe or discuss their problems in their own voice. This has motivated several researchers to

[1]http://reddit.com/
[2]https://www.reddit.com/r/SuicideWatch/
[3]https://www.reddit.com/r/Depression/
[4]https://www.reddit.com/r/Bipolar/

probe the different aspects of support online, like language [CK17], anonymity in social support [DD14] or risk factors [Gko17]. Recent studies have shown promising results in modelling and measuring signals and patterns in reddit communities related to mental health. For instance, statistical relations of mental health and depression communities with suicidal ideation have been studied [DD14, DCKD^{+}16]. The authors explored linguistic and social characteristics that evaluate user's propensity to suicidal ideation. Approaches to classify reddit posts as related to certain mental health conditions have also been successfully developed, showing that there are certain characteristics specific to mental health-related topics in posts that can be automatically captured[Gko17]. Furthermore, in a study focused on reddit posts related to anxiety, depression and post-traumatic stress disorder, the authors show that these online communities exhibit themes of supportive nature, e.g. gratitude for receiving emotional support[PCC18]. Positive effects in participation in such fora have also been shown by improvements in members' written communication[PC18]. The supportive nature of comments in the SuicideWatch forum has also been studied by automatic identification and classification of helpful comments with promising results[KRMH^{+}16].

One of the most challenging, and devastating, global mental health concerns is suicide. Suicidal behaviour includes any thoughts, plans or acts someone makes towards ending their life. In health care services, preventing death by suicide is a priority, but accurately predicting whether or not someone is at risk of committing suicide is difficult. Moreover, a large proportion of deaths by suicide occur in populations that have never been seen by health service providers. Several online platforms are used for expressing suicidal thoughts and reaching out for support. On Reddit, the subreddit *SuicideWatch* currently[5] has almost 94k subscribers, and is a lightly moderated forum that is intended to offer peer support for people at risk of, or are worried about others', suicidal behaviour. The moderators take the intent of peer support seriously, and are governed by guidelines that prohibits false promises, abuse, tough love and other clinically frowned upon methods of conversations[6] This setup gives us an ideal petri-dish to measure what constitutes a supportive conversation. More specifically, owing to the results from Chapter3, we would like to develop methods that would allow us to

[5] As of 27th June 2018
[6] https://bbc.in/24rJYQH

understand the macroscopic and mesoscopic signatures of support. Formulating this problem needs to follow the methodological framework discussed in Chapter 1, whereby we collect the necessary data and device appropriate abstraction/s, to quantify how the macroscopic and mesoscopic signatures of support in these conversations manifests. These signatures could be captured using different metrics that signify a particular behaviour of interaction.

4.1.1 Research questions

Most previous studies have aimed at studying the *content* of posts and their characteristics in relation to other posts. One important aspect of online communities is its supportive *function* — users turn to these platforms not only to express their thoughts and concerns, but also to receive support (instrumental or perceived) from the community. In the previous chapter, we looked at how this received support, manifests in the network structure in the short term and long term users. What remains to be explored is how these supportive conversations themselves look on a macro as well as local (meso) scale. More specifically, in this chapter we would like to arrive at answers to the following two research question, discussed in Chapter 1

RQ3 *Are there any macroscopic signatures of supportive conversations?*
RQ4 *Are there any mesoscopic signatures of supportive conversations?*

To model the network topology in an online community, we represent each conversation happening over these forums using graph-based abstractions (users and replies) as described in Section 4.3. To measure global structure of these conversations, we user network topological metrics such as centrality: which measures importance of nodes in a network in terms of relaying information, branching factor: which measures how a conversation fans our over time, return distance: which measures how soon do users return back to the conversation and symmetric edges: which measures reciprocity of users in a conversation. To measure measure local interactions, we measure inter response times: which measure urgency of

Terminology	stands for
RP	Root post which begins a new thread on a subreddit
OP	Original poster who posts the Root post for a thread
SW	The suicide watch Subreddit
FP	Front page of Reddit.

Table 4.1 Notations and Terms.

response to a message, semantic alignment between messages and local interaction motifs known as Triadic motifs : which gives an idea about how distinctive are interactions between subgroups of users.

4.2 Data

Reddit is a platform where a user can create a post on a subreddit, and other reddit users can interact by posting at different levels of the thread or by up or down voting posts. We analyzed root posts in the SuicideWatch subreddit (SW)[7], building on the work of Gkotsis et al. [Gko17]. We crawled SW to get hierarchical threaded conversations, by iteratively pursuing each conversation at progressively deeper levels until the whole thread had been obtained. [8]. This resulted in a dataset of over 50k SW threads totaling around 500k individual posts.

To provide a baseline against which to compare the hypothesized supportive nature of the SW sub-reddit, we acquired 50k baseline posts from any other subreddit popular enough to land on the frontpage.

We crawl the Frontpage posts for 2 weeks accumulating over 50 thousand reddit threads in the process. Comparing the suicide watch threads with the baseline, gives us confidence over our network metrics, and any claims made there forth. The two conversation datasets from r/SuicideWatch and Frontpage are very similar in terms of pure macro statistics. E.g. the median amount of responses for a Suicide watch thread were 6 and for baseline Frontpage posts were 8. The median conversation depth (the depth of the hierarchy of the responses)

[7]https://www.reddit.com/r/SuicideWatch/

[8]The code to crawl reddit for threads can be found at *https://github.com/sagarjoglekar/redditTools*

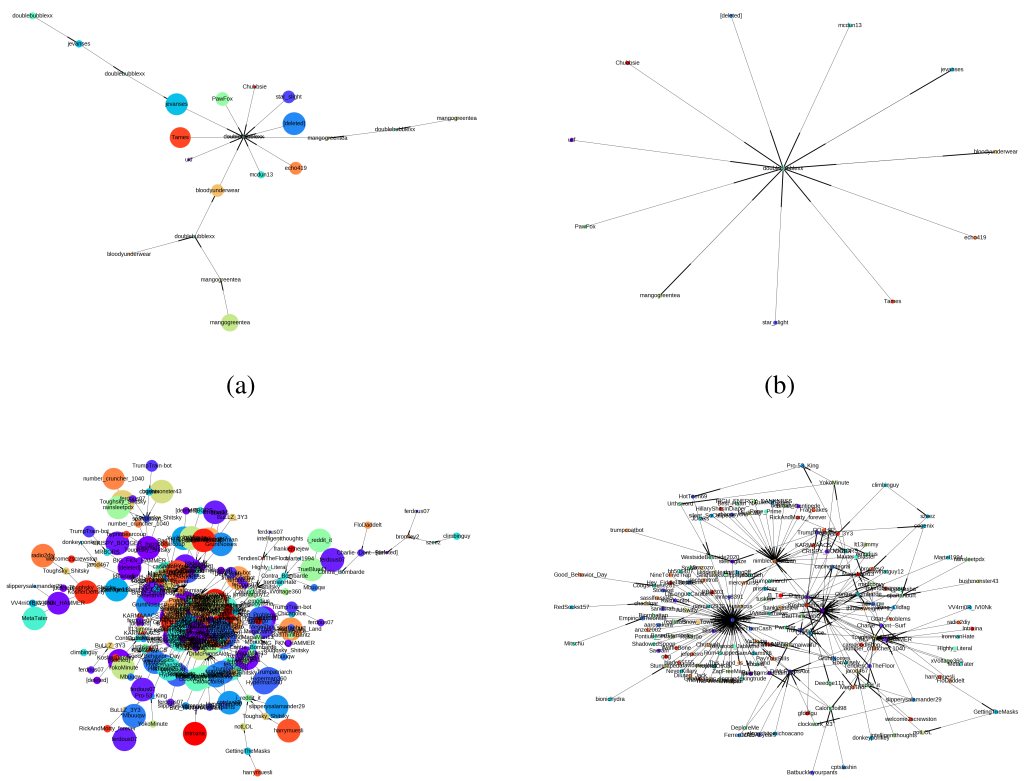
is 2 for all the datasets. The median amount of unique authors participating in a post are 2. Owing to the long tailed nature of the datasets, we perform our analysis on threads which have been conversed on at least 5 times. Which means we consider graphs which have more than 5 posts on the root post. This results in our datasets to shrink to 25k threads for suicide watch and 30k threads for Baseline. We further clean the dataset, by removing threads where the root author has deleted their user account, which is a common practice to preserve anonymity in more controversial posts. The resulting dataset has 20k threads in Suicide watch and 23k threads in the baseline.

4.3 Abstractions

To understand the dynamics of supportive conversations, we first need to formalize the abstraction of networked conversations as well as the content posted in these conversations. In case of forum based platforms where users interact in a nested dialogue fashion, and original poster or *OP* posts a start of a thread. This thread is then open for comments by all the community users. In case of Reddit, such a community is called a Subreddit, which is a moderated collection of users who subscribe to it. These users may post new threads onto the subreddit as far as the post follows the subreddit rules. Enforcement of these rules is the responsibility of the moderators. The user who starts a thread is called the Original Poster or **OP** and the headlining post which the *OP* begins with is called the Root Post or *RP*.

Reply Graphs

The first abstraction mimics directly the structure of conversation threads on Reddit. These abstractions are called Reply Graphs. We formulate a reply graph $R\{P,E,W\}$ as a thread of multi-layered posts in a thread in response to the root post *RP* in the sub-reddit. Each graph *R* consists of posts $P_i, P_j, i, j \in N$, where N+1 is the total number of responses in the thread and edges E_{ij} such that and edge E_{ij} exists iff post P_i was in response to post P_j in the hierarchy of responses. The weight of the edge E_{ij} is found by calculating the cosine similarity between semantic vector V_i for post P_i and the semantic vector V_j of post P_j. This abstraction works



(a) (b)

(c) (d)

Fig. 4.1 Example User graphs and their corresponding Reply graphs, Figure 4.1a shows a random thread from the Suicide Watch sub-reddit along with the preserved structure of posts in the thread and 4.1b shows the corresponding user graph that arises by capturing the user to user interaction in the thread from the reply graph. In comparison we have user graph Fig 4.1d and its corresponding reply graph Fig 4.1c from one of the Front page threads

well in modeling the conversational nature of these forums. For convenience of the reader, we present a couple of example pairs from SW and Frontpage baseline datasets in Figure 4.1

User interaction Graphs

In this method, we represent each thread as a directed graph $G\{V,E,W\}$ where V is the set of all users participating in a particular thread and E are the directed edges which correspond to interactions between two users $V_i, V_j \in V$. The weight of each directed edge E_{ij} corresponds to the average of all the edge weights between $V_i, V_j \in V$ in the corresponding reply graph $R\{P,E,W\}$ as described above. This means that each reply graph is then mapped to a User graphs where the nodes are users rather than posts. Another salient distinction between the two abstractions is that reply graphs resemble an n-ary tree and user graphs are directed cyclic graphs.

4.3.1 Network characteristics

Figure 4.2a shows the distribution of maximum depths across all Reply graphs for SW and Baseline subreddits. The SW threads depths have a median depth of 2 and mean of 4 compared to median depth of 2 for BL and a mean of 2.5. This shows that statistically the depths of Suicide watch and baseline graphs are quite similar.

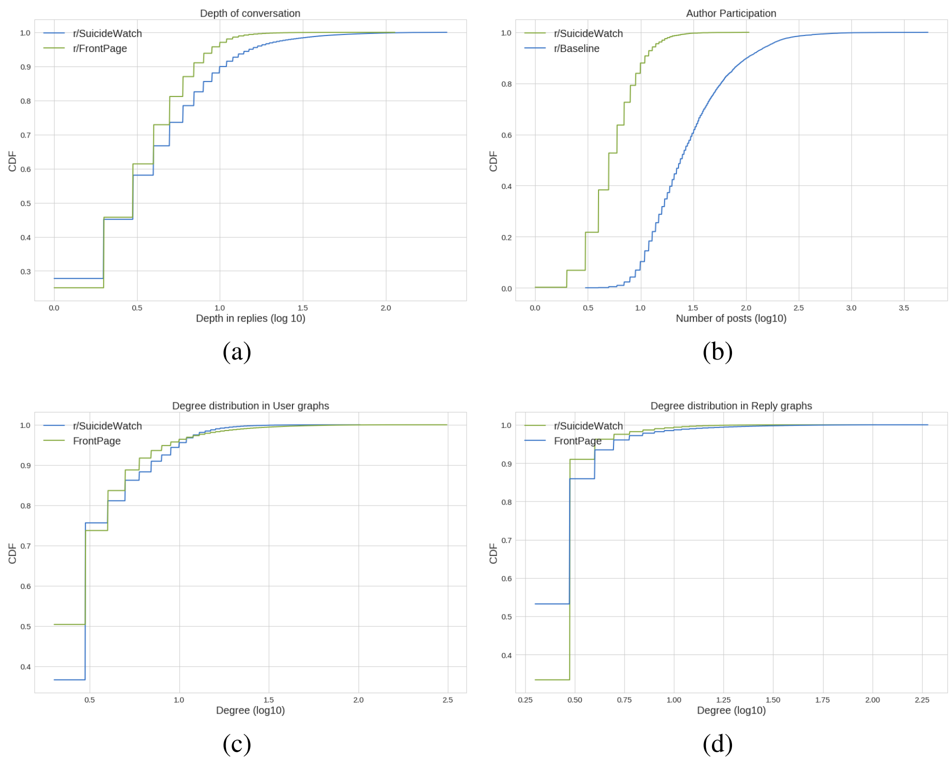


Fig. 4.2 *Fig 4.2a shows the distribution of maximum depths of Reply Graphs for Subreddit r/SuicideWatch and the baseline Frontpage conversations. Fig 4.2b shows the distribution of unique authors per thread in the two datasets. Fig 4.2d shows Distribution of degrees for Reply Graphs, r/SuicideWatch and FrontPage. Fig 4.2c shows the degree distributions for the reply graphs*

4.4 Metrics

Once you are able to develop abstractions from the data, it becomes possible to develop metrics on these abstractions so as to derive insights about the mechanics of support. These metrics are then used to validate structural differences between supportive conversations and generic casual conversations from our baseline set, and come up with a theory for links between supportive conversations and the structure of the conversation, if we find any.

Semantic Alignment

We then extract the word embedding vectors for each post using Doc2Vec[LM14], which extends the word embeddings to represent a whole document or paragraph. We extract these embedding vectors for each post in R^N for all the posts across the complete hierarchy of threads. We then quantify the edge weights of each interaction amongst the reply tree as the cosine distance between the response post and the hierarchically higher post, to which the responder has posted to. This captures the semantic alignment between the hierarchically adjacent responses. More formally, if user V_i has responded with post P_i to a post P_j written by user V_j, we extract the word embedding in R^N for both posts P_i and P_j. If these embeddings are represented by N dimensional vectors ψ_i^N and ψ_j^N then the edge weight for the edge E_{ij} in the corresponding reply graph would be

$$W_{ij} = \frac{\psi_i^N . \psi_j^N}{\|\psi_i^N\|_2 \|\psi_j^N\|_2}$$

This metric standardizes all edge weights between 0.0 and 1.0, 1.0 implying that the posts P_i and P_j are most aligned, and 0.0 implying the post have least semantic similarity. This metric abstracts out the content of the post, in terms of semantics which can then be used as edge weights in the graph abstractions.

We use a popular word embedding method called *Word2Vec* [MSC$^+$13] which learns representations of a set of words from a corpus of text, which in our case is the text from Suicide Watch and baseline fora. These representations can be used to extract text embedding vectors for each post which belong to a N dimensional space R^N. These vectors are tested for their alignment using cosine distance in R^N, which from literature is shown to correspond to the semantic similarity in the textual space. This method is quite popular and used in community based question answering[MN16], Medical semantic similarity [DVZK$^+$14] and other medical informatics applications[ZYW17]. We first train two independent word2vec models on the Suicide watch and Front page post corpora. We then extract the word embedding vectors for each post using Doc2Vec[LM14], which extends the word embeddings to represent a whole document or paragraph. We extract these embedding vectors for each post in R^N for all the posts across the complete hierarchy of threads. We then quantify the edge weights of each interaction amongst the reply tree as the cosine distance between the response post and the hierarchically higher post, to which the responder has posted to. This captures the semantic alignment between the hierarchically adjacent responses. More formally, if user V_i has responded with post P_i to a post P_j by user V_j, the edge weight of of edge E_{ij} in the corresponding reply graph would be

$$W_{ij} = \frac{V_i.V_j}{||V_i||_2||V_j||_2}$$

This metric standardizes all edge weights between 0.0 and 1.0, 1.0 implying that the posts P_i and P_j are most aligned, and 0.0 implying the post have least semantic similarity. This metric abstracts out the content of the post, in terms of semantics which can then be used as edge weights in the graph abstractions.

Centrality

For this metric we use the User Graphs. Node centrality is a metric that measures how central a node is in a network. It directly reflects the importance of the node when it comes to membership of the shortest connecting paths between all the nodes in the graph. More formally, we use betweenness centrality of a node which is defined as

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where $\sigma_{st}(v)$ is the total number of shortest paths from node s to node t and σ_{st} is the number of those paths that pass through v. To understand whether the thread starters (*OP*) have a special place in the network, we evaluate both centrality of the node corresponding to the *OP*, as well as median centrality across all the nodes in a user graph.

Symmetrical users

We define a symmetric user and a symmetric edges for user graphs. For a user V_i in the user graph $G\{V,E,W\}$ as described in Section , a symmetric user is a user who interacts with any user V_o or the *OP* and receives a response back from that user or the *OP*. We find the fraction

$$U_{sym} = \frac{total\ number\ of\ symmetric\ users}{Total\ users\ in\ a\ thread}$$

Urgency

To understand the urgency in how Suicide watch subreddit users responds to the *OP* and each other, compared to the baseline threads on the Frontpage, we calculate differences between the posting times between consecutive response messages in a reply graph. We then compute the median response times per thread, for posts in response to any *OP* authored posts and in general across all other post responses.

Branching Factor

Branching factor is a quantity that reflects the fan out of a conversation as it evolves. To measure this phenomena, we use the reply graphs, which resemble a n-ary tree, to evaluate the branching factor. The branching factor is formally described as

$$\tau = \frac{1}{|D|}\sum_{d\in D}\frac{1}{|N_d|}\sum_{n\in N_d} InDeg(n)$$

4.5 Macroscopic analysis: at conversational level

Through our analysis we find several discriminatory factors among Suicide watch conversations and generic front page conversation. We show that some of these factors are archetypal of suicide watch conversations such that they are over-expressed in suicide watch conversations to a very high degree. We also show that certain properties of these conversations can be backed by sociological theories of real life support conversations.

4.5.1 Peculiarity of threads of Support

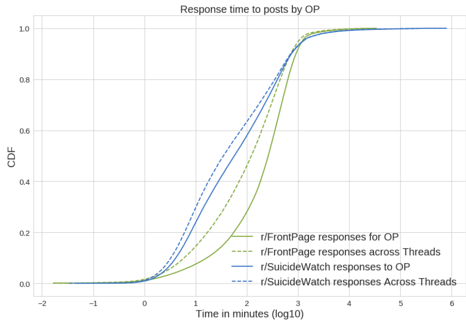
We begin by characterizing the two networked abstractions, namely Reply Graphs and Interaction graphs as described in Section 4.3. We do so by first comparing these two abstractions with a baseline control conversation threads using certain macroscopic network properties.

Responses to *OP* are very urgent in supportive setting

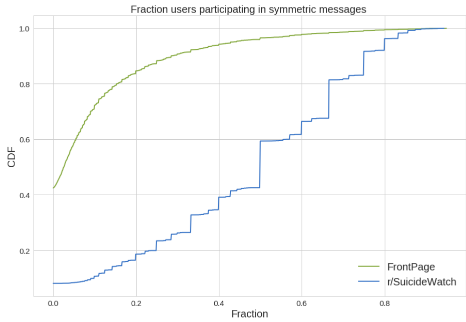
Understanding the inter message times can act as a good proxy for the urgency in a conversation. To understand how Suicide watch subreddit users responds to a *OP* and other users as compared to other sub-reddit threads on the frontpage, we calculate differences between the posting times between consecutive messages in a reply graph. Figure 4.3a shows comparison using CDFs of inter-message response times for SW and FP threads. It can be seen that SW *OP* are responded with the highest urgency amongst the 4, especially compared to either the *OP* or any other users or sub-reddits.

Interactions on suicide watch forums are statistically more symmetric

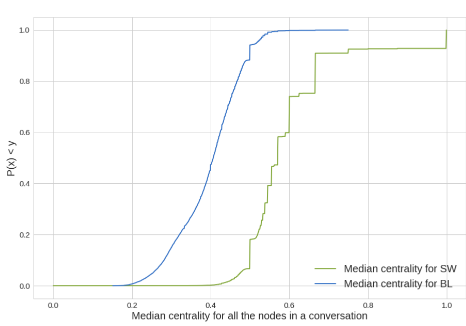
Despite signs of urgency and engagement, we ask the question: what percentage of conversations happening on these subreddits are symmetric in nature ? For this The median value for U_{sym} for SW is 20% where as for AS is 0%. This shows that SW subreddit engages is a lot more symmetric conversation that the baseline threads. If we define a set of users who engage in



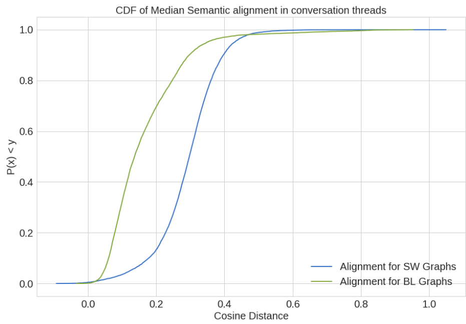
(a)



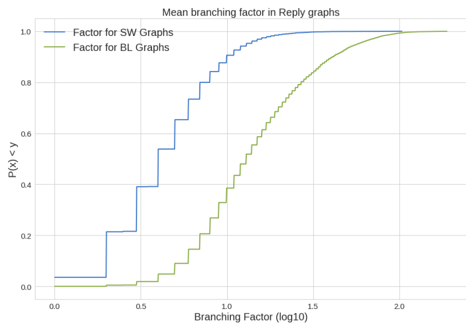
(b)



(c)



(d)



(e)

Fig. 4.3 Panel shows CDFs of different network metrics. Fig.4.3a shows the response time distributions, Fig.4.3b shows symmetrically engaged users, Fig.4.3d shows topical similarities across posts and 4.3e shows the branching factors of reply graphs.

symmetric activity with the *OP* , it would be worth while to investigate how much of the total message activity on the thread is carried out by these set of symmetric users . To calculate this we find the fraction of messages on each thread written as part of this symmetric conversation. Figure 4.3b shows the trend. It can be see that SW threads contain a higher prevalence of symmetric message exchanges compared to the baseline Frontpage threads. This shows a higher engagement from the *OP*s side when participating in a supportive conversations

OP is the most central user in supportive conversations

To understand how embedded is the *OP* in a conversation thread, we compare the betweenness centralities of *OPs* in the *SW* dataset with the baseline FP dataset. Betweenness centrality is a good proxy of understanding how closely linked is a node with the rest of the network. When we calculate this metric for the user graphs we see that Suicide watch *OP*s tend to have highest centralities compared to generic FP threads moth in terms of *OP* centrality as well as median centrality across all the users. The high centrality of *OP*s in *SW* conversations implies a high level of embedded-ness as well as a *OP* centric approach by other participants in the conversation. The Figure 4.3c shows the Empirical CDFs of centralities.

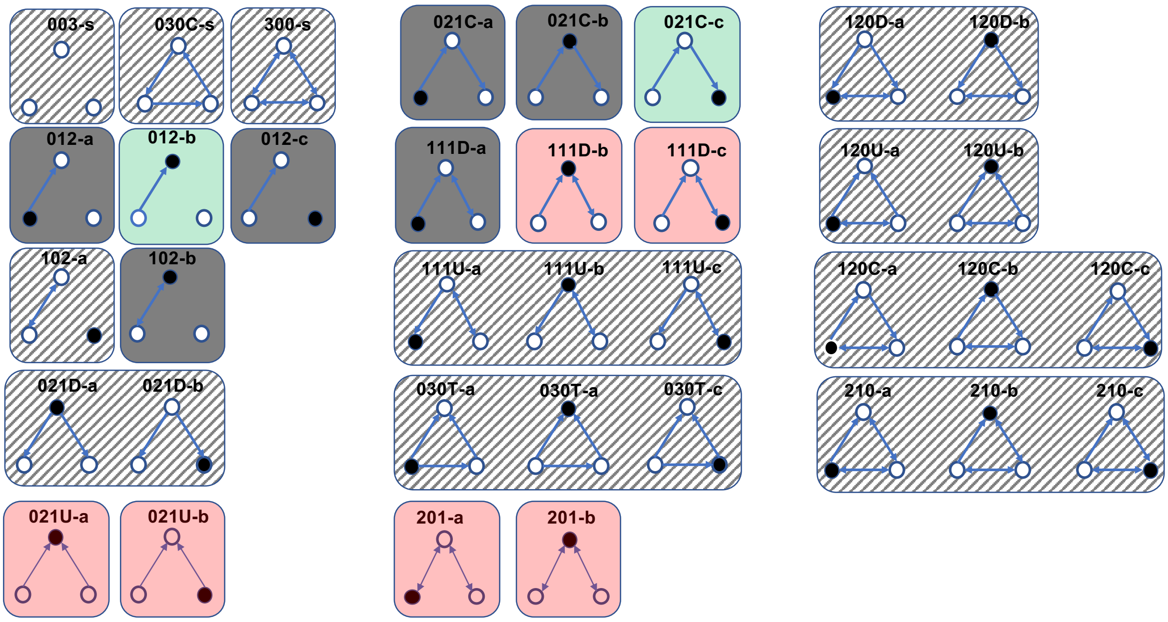
Supportive responses are semantically aligned, more so when they are in response to the *OP*

We measure semantic alignment based on word embeddings of the source post and the reply post, at every edge of the reply graph. The detailed method of extracting semantic alignment along a post and its response is described in Section 4.4. Extracting such similarity metrics, we compare the trend in response text being in semantic alignment with the parent text in the reply graphs.

Supportive conversations branch out considerably less compared to baseline

Branching in a conversation thread could be either a sign of digression or a sign interestingness resulting in more people joining in. To measure this phenomena, we use the reply graphs, that mimic the conversation structure of the threads. By using the method described in Section 4.4, we found that Suicide watch threads, tend to branch less as compared to our baseline conversations. This implies that suicide watch threads tend to remain on topic and more often than not, a one-on-one conversation. Albeit many such dialogues may emerge with many participants, and hence that explains the high centrality of the *OP* in all user interaction graphs. If the participants on a thread seldom interact amongst themselves, the corresponding interaction graph will have the *OP* as the most central node.

4.6 Mesoscopic analysis: Anchored triadic motifs.



(a)

Fig. 4.4 Figure 4.4a shows the 36 different types of Anchored Triadic motifs that are looked for in the user graph data. The motifs with **green boxes** are **over expressed in the baseline dataset** by significant amount. The motifs with **red boxes** are **over-expressed in the Suicide watch dataset** by significant amount. The motifs with **grey boxes** are present in significant numbers in both datasets, but neither over nor under expressed in any datasets based on their Z scores. The motifs in **grey hatched boxes** are very rare in both the baseline and suicide watch datasets, with less than 5 mean occurrences per graph per bin.

Network motifs are local sub-networks between 2 or 3 nodes. Such local patterns are highly useful in quantifying local interactions and the resulting macro structure of the network[MSOI^{+}02]. They have been used in a variety of applications and networks, from economics [ZSSH14] to cellular protein-protein interaction networks [YLSK^{+}04]. These local interaction patterns have been studied before, and have been fundamental in the study of social structural processes[Fau07]. They help social scientists quantify the type of hierarchies in the social network[Dav67, DL67]. But in the setup of a typical support community network like r/SuicideWatch, the conversation shows clear distinction between the users who respond to a call for help and the user/s who are asking for help. In the case of Reddit we define the former as the *OP* who starts the conversation thread. For this reason, we changed the definition of the triadic motifs to accommodate the role of a node around which the motif is developed in the due course of the conversation. In conventional literature, the local interactions are measured in terms of census of 16 triadic motifs[Fau07], which cover all possible patterns of non-isomorphic graphs, or graphs which cannot be mapped or morphed into each other. In this method, there is no special treatment to any node, and positional parity of all nodes is treated equally. This means, role of a node in the conversation cannot be incorporated in this analysis. For this reason, we build on top of this methodology but by introducing anchors, or nodes with special importance. This is because, we would like to know how a graph evolves around the person in distress or the *OP* in our case. A symmetric link with the *OP* implies that the *OP* is part of a to and fro communication, however a triad with a symmetric link between two non-*OP* nodes implies that the conversation does not involve the *OP*. By fixing a role for a node in a motif, each of the 16 triadic motifs as seen and developed in the field[Fau07, HL77], can be unravelled into 36 sub-variants of these motifs by varying the anchored node, as seen in Fig 4.4. Each sub-variant is different from the other from the perspective of the anchored node. The resulting motifs to measure grow from 16 to 36, since some anchored motifs are iso-morphic, which means they look identical to some other motif from the perspective of the anchor. These motifs are then counted using custom tweaked census methods based on Bataglej et.al's work[BM01]. Each motif as seen in Figure 4.4a is named using a particular recipe. The first three letters, follow a M-A-N pattern which signifies the number of "**M**utual" , "**A**symmetric" or "**N**ull" edges present in that particular triad. E.g. the motif 030 has 0-Mutual(bi-directional), 3-Asymmetric(unidirectional) and 0-Null(disconnected) edges. There are some motifs with an added letter modifier attached, to signify the flow of information in the triad. The naming is done in accordance with Holland et. al's work[HL71].

To methodically understand the over or under expression of these anchored triadic motifs in the suicide watch community, we use the user interaction graphs for the Front page baseline posts as a control group. We analyse 10,000 user interaction graphs each for the baseline and the suicide watch datasets. We progressively select graphs with variable sizes i.e. number of users present in the interaction graphs. We do so by binning each dataset in ranges of graph sizes in increments of 5 nodes per bin. The resulting graphs would fall in the following 7 bins, with increasing number of nodes present : 1 - 5 , 6 - 10 , 11 - 15 , 16 - 20 , 21 - 25, 26 - 30, 31 - 35 and 36 - 40. We stop sampling above 40 nodes per graph, as the population of conversations that contain more than 40 unique users participating is extremely scarce for both the datasets. We then do the census for the 36 unrolled motifs for each of these bins, for both the datasets. Once the census is done, we calculate Z_{scores} for the Suicide watch conversations, using Baseline conversations as the null model, to understand over or under expressions of certain motifs. For any given bin B_I, let there be k baseline graphs that fall in that particular bin noted by G_{BL} , and let there be n suicide watch graphs falling in the same bin signified by G_{SW}. For such a setup, let M_{BL} signify a vector of k elements, where each element is the total number of occurrences of particular motif m in each graph from the set G_{BL}. With this sample of graphs as the null model, the mean would be $\mu_{null} = \frac{1}{k}\sum_{m_i \in M_{BL}} m_i$, where m_i is the count of motif m in the i^{th} graph in G_{BL}. The standard deviation σ_{null} is $std(M_{BL})$ Once we have the null model parameters for the baseline graphs for a particular bin, we calculate Z_{scores} for all the graphs from suicide watch samples G_{SW} from the same corresponding bin as random variable Z where the i^{th} element is the Z_{score} for graph i in G_{SW}. The score is calculated by the following formula

$$Z_i = \frac{m_i^{sw} - \mu_{null}}{\sigma_{null}}$$

where m_i^{sw} is the total number of occurrences for motif m in the i^{th} graph for the suicide watch samples. We then plot the mean of random variable Z and the standard error in Z to understand the over/under expression trends.

4.6.1 Patterns in local interactions

It is often useful to express large interaction graphs, as the sum of local interactions between two or three nodes at a time. This method is quite prevalent in the Social sciences, for studying social structures by looking at local interaction between agents[Fau07]. Such analysis is quite useful in expressing local structures in the graphs and has been used in several network analysis works[WLH14, SM15]. For this reason we conduct a census of the 36 Anchored triadic motifs (described in Section 4.6) across all the selected graphs. From the amount of over or under expression of the network motifs, researchers have made inferences about the nature of local interaction. They do so by comparing the amount of density of each triad in a real network as against the expected quantity in a null model based on the number of edges[Fau07]. We perform binning of user graphs as described before in Section 4.6, and perform over or under expression analysis in comparison with the baseline null model, using Z-scores of the motif occurrences.

We find that anchored motif variants **021U-a, 021U-b, 111D-b, 111D-c, 201-a and 201-b** are significantly over expressed in suicide watch conversations across all sizes of graphs as seen from figures 4.5a,4.5b,4.5c,4.5d, 4.5e,4.5f. Similarly anchored motif variants **012-b and 021C-c** are significantly over expressed in the null model(baseline) graphs across all sizes.

From previous studies on triadic structure, it was inferred that transitive triads are naturally more common than expected in social structures of human social networks. Interestingly, our analysis shows that transitive triads are rarer in Suicide watch, as against the baseline conversations. But in the defence of previous studies, the networks we study are of human conversations and not of human social ties. Also these conversations are happening with an intent of providing support, which makes one user (the one in distress) the centre of conversation. These conversational preferences make the macro and meso level structural signatures peculiar.

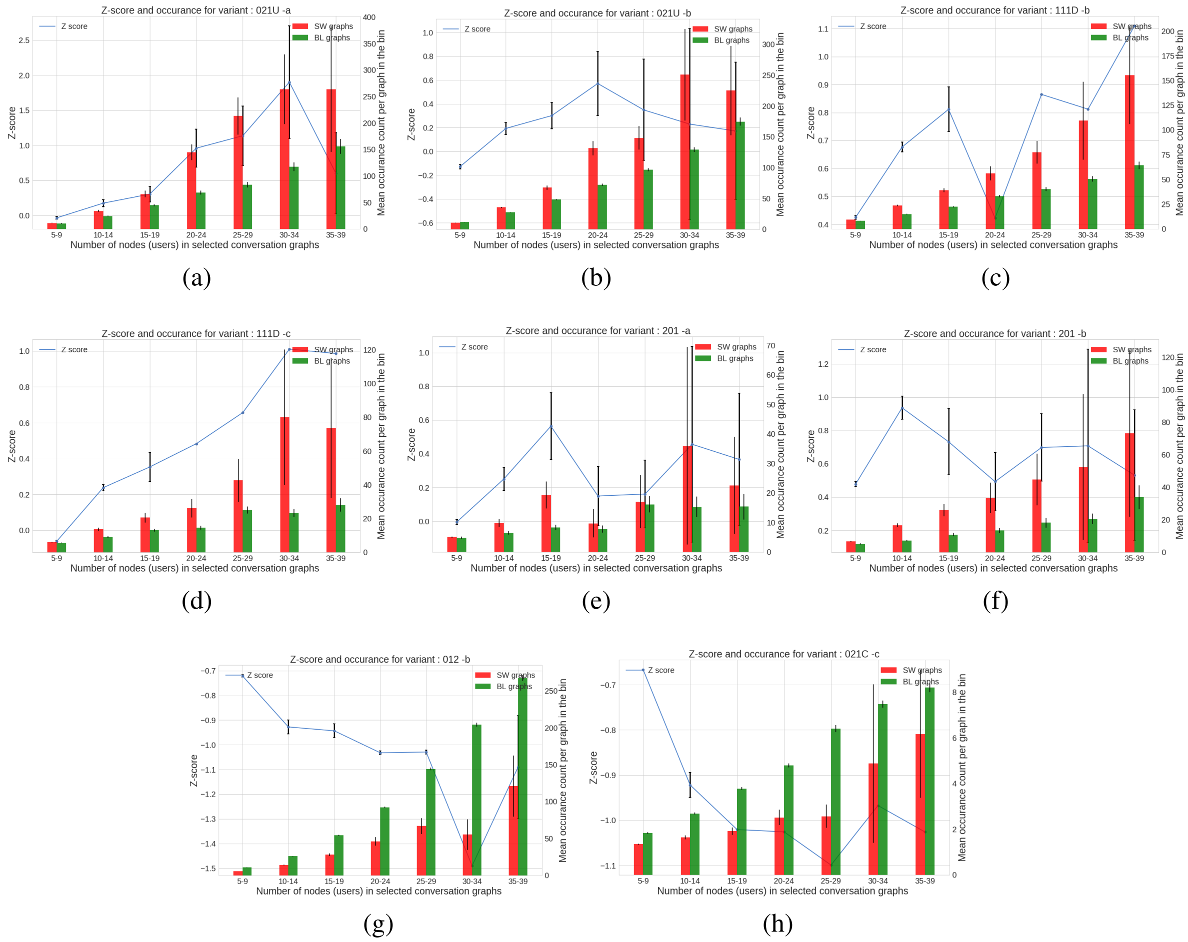
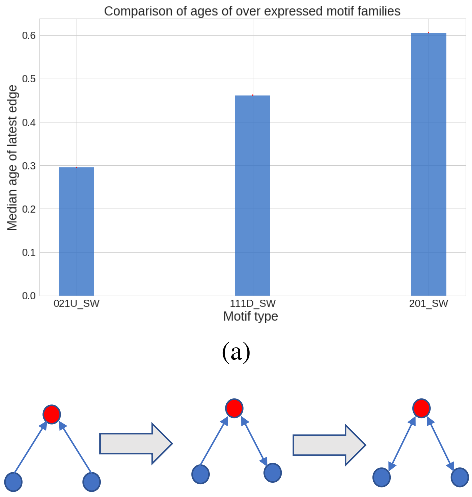


Fig. 4.5 This panel shows the Z-scores of the abundant and over expressed Anchored motifs found in both datasets. There are variants which are over expressed in Suicide watch (Like figures 4.5a,4.5b,4.5c,4.5d, 4.5e,4.5f), which form the top six plots. The bottom 2 plots show motifs which are over expressed in the Baseline over Suicide watch, as their Z-scores are negative. Figure 4.5g and Figure 4.5h.

4.7 Conclusions and Outlook

These patterns in local interactions indicate that conversations in suicide watch tend to be more *OP* centric, with non-transitive dialogues between users who respond to a call for help. More so, the *OP* tends to be highly central in the conversation as well as part of several mutual interactions. These behaviours are highly particular to r/Suicidewatch when compared with the baseline conversations. Further investigation of local structures in graphs show that there is a radically different way the conversation graphs evolve. The difference



(a)

(b)

Fig. 4.6 *Figure 4.6a shows the plot for mean motif life times as a fraction of the entire thread life. The lifetime is measured as the time at which the latest edge of motif is formed. The Figure 4.6b shows a toy example of what sort of dialogue exchange could lead to the motif lifetime distributions seen in Figure 4.6a*

is not only between supportive conversation graphs and generic conversations, but also overall in supportive conversations and any social graphs. Several studies have looked at the social-tie structures and shown certain triadic motifs to be important in maintaining social hierarchy [SM15, SM12, HL76, HL71]. But these expected structures are not the ones prevalent in supportive conversations. Transitive triads like *030T* and *210C* are shown to be overexpressed in ties that show social hierarchies. This points to a very different approach towards interactions with distressed users, and may have great implications on detection of toxic behaviour and enabling of helpful online conversations. Further work needs to be done in modelling these conversations. With the advances in machine learning, both the language and the network structure could be used in compliment to develop unique embeddings for support. A true inter disciplinary aim is to enable use of such frameworks, to understand at risk individuals online and offline.

There are other relevant triads are not even active in supportive conversations as seen from Figure A.0 (Discussed in Appendix 1). The most interesting ones are the motifs that exhibit triadic closures, which have been shown to be fundamental for social capital in networks [BDIF14, Gra77, JK19]. The absence of these motifs in either Suicide watch or baseline graph structures indicates that the conversational graphs online evolve differently to social graphs. This phenomenon needs to be explored further.

In conclusion, quantifying perceived social support is an ongoing effort. But my work has shown that there are computable methods, which could capture the mechanics of support in online conversations.

4.7.1 Implications

Online spaces are used by a diverse set of users. One of the main challenges in front of open online spaces like Twitter and Facebook is to make them safe and toxicity free for users. This however is in direct contradiction with the free and open philosophy that underpins the creation of the internet. At such a juncture, it is worth investigating a setup where these platforms offer a differential policy towards toxic or non mainstream behaviour online. The main challenge in allowing differential behaviour policing, is to understand where to draw the line. My work sheds some light towards finding specific interactions online which are the best candidates for policing for toxic behaviour. By quantifying the structures of peer support in online spaces, we can enable public platforms to be allow serendipitous supportive interactions to happen, without clamping down on the open nature.

Chapter 5

From Crowds: Real world perception of beauty

Beauty is nothing other than the promise of happiness – Stendhal, *On Love*

In part one, I delved deeper into the problem of quantification of perceived social support through online communities. This work showed that there are quantifiable signatures of social support in the structure and content of online communities, and these can be exploited to build models of supportive and safe interactions online, which could directly benefit persons in distress. Capturing the signatures of a subjective quantity, like the perception of social support, through the analysis of online conversations has implications on how we design better online spaces in the future. This feat is achieved through investigating networked, interacting users in communities. But the question to ask is, can unconnected agents, in large enough quantities(crowds) be use to extract signatures of perceptions.

In this spirit , in the second part of my thesis, I extend the motif of perception driven design of spaces, to the offline world. The core idea is about asking a similar question, that is: **"Can we quantify the signatures of subjective perception of the real world through crowd opinions?"**.

It has been shown through multiple studies, that the physical spaces that we use have measurable effects on our health [MVG+06, LM11], our outlook towards exercise [TGL+14], health of seniors [TNW02] and general all round well being [GTMM+15, NPK13]. More importantly, all these studies point to the value of aesthetically pleasing spaces for an overall liveable city. This intuition motivates the choice of quantifying the aesthetic perception of real spaces. More so it motivates the question: Can crowd's opinion help capture a subjective quantity like the aesthetic? However, the question about whether an urban space is considered beautiful or not is highly subjective. To that end, we need to refine the hypothesis that we are aiming to prove. The perception of aesthetic in urban spaces are affected by location, culture, background of the crowd member and several other subjective attributes. This is why, I aim at quantifying what aspects of an urban scene "on an average" are found aesthetically pleasing to the crowds.

Research has shown that there are specific categories of urban elements that are universally considered beautiful e.g. greenery, small streets, memorable spaces, open skies [Ale77, QOC14, SSH13]. These elements are those that contribute to the creation of what the urban sociologist Jane Jacobs called 'urban vitality' [Jac61]. The idea of this work is to test the quantifiable and predictive nature of these elements. In this section of my dissertation, I show that this is feasible using cutting edge deep learning and computer vision techniques and some metric design which associates meaning to the patterns in data.I use google streetview images as the source of data for photographs of urban scenes. I use crowd opinion to capture the predictive motifs of urban beauty. And I use literature drive metrics to explain and quantify the urban perception of beauty. This follows that in this chapter, I would like to answer the RQ5 of my dissertation.

RQ5 *Can crowdsourcing and machine learning help us quantify how humans perceive beauty in urban settings?*

5.1 Related Work

The problem of designing better cities, has been a on going obsession for various fields. A lot of work was done in the past in just understanding the concept of good , liveable cities. One of the most prominent figure in the campaign to re-vitalize our cities was Jane Jacobs in the late 50s. Jacobs's seminal work on urban vitality [Jac61] discusses the idea of how a design of a city might be the driving reason behind how urban vitality thrives. Christopher Alexander was another such prominent voice who in his book undertook a cataloguing exercise of typical "patterns" of good urban design [Ale77]. This effort showed that certain patterns of placements of roads, trees, walkways, parks and areas of social interactions are highly crucial in promoting a thriving social environment. In the fields of psychology, environmental design and behavioral sciences, research has studied the relationship between urban aesthetics [RAS00] and a variety of objective measures (e.g., scene complexity [KKW72], presence of nature [KK89]) and subjective ones (e.g., people's affective responses [Ulr83]). As mentioned before, the relation between greenery and the design of spaces around us has been linked with measurable effects on our health [MVG+06, LM11], our outlook towards exercise [TGL+14], health of seniors [TNW02] and general all round well being [GTMM+15, NPK13].

With the premise of all this, it is worth exploring the literature to understand how different fields are working towards using this relation between the humans and the spaces they occupy.
Ground truth of urban perceptions. So far, the most detailed studies of perceptions of urban environments and their visual appearance have relied on personal interviews and the observation of city streets: for example, some researchers relied on annotations of video recordings by experts [SR04], while others have used participant ratings of simulated (rather than existing) street scenes [LH12]. The Web has recently been used to survey a large number of individuals. Place Pulse is a website that asks a series of binary perception questions (such as 'Which place looks safer [between the two]?') across a large number of geo-tagged images [SSH13]. In a similar way, Quercia *et al.* collected pairwise judgments about the extent to which urban scenes are considered quiet, beautiful and happy [QOC14] to then recommend pleasant paths in the city [QSA14]. They were then able to analyze the scenes together with their ratings using image-processing tools, and found that the amount of greenery in any given scene was associated with all three attributes and that cars and fortress-like buildings were associated with sadness. Taken all together, their results pointed in the same direction: urban elements that hinder social interactions were undesirable, while elements that increase interactions were the ones that should be integrated by urban planners

to retrofit cities for happiness. Urban perceptions translate in concrete outcomes. Based on 3.3k self-reported survey responses, Ball et al. [BBLO01] found that urban scenes with positive aesthetics properties not only are visually pleasurable but also promote walkability. Similar findings were obtained by Giles et al. [GCBK+05].

Deep learning and the city. Computer vision techniques have increasingly become more sophisticated. Deep learning techniques, in particular, have been recently used to accurately predict urban beauty [DNP+16, SPM17], urban change [NKR+17], and even crime [DNVZ+16a, AERA14]. Recent works have also showed the utility of deep learning techniques in predicting house prices from urban frontages [LSSGR18], and from a combination of satellite data and street view images [LPR18].

To sum up, a lot of work has gone into collecting ground truth data about how people tend to perceive urban spaces, and into building accurate predictions models of urban qualities. This trove of human annotated ground truths about urban spaces is vital in understanding human perception at scale. In this chapter we would look at a way to transfer the collective perception of humans in to a machine learning model. Doing so we would validate whether machine learning models can actually capture the subjective perceptions of people.

5.2 The Data

To begin with, we need highly curated training data with labels that reflect the crowds consensus on beauty. Unfortunately, it is difficult to get data where thee is a consensus on of the crowds on an absolute value of beauty. However there are datasets available where crowds have voted on pairwise comparisons of urban images for their relative beauty. The two most prominent examples are from MIT media lab [DNP+16] called the place pulse and from Bell labs [QPAC13].

We start with the Place Pulse dataset that contains 100k Google Street Views across 56 cities around the world [DNP+16]. The voting on these scenes is taken in the form of a gamified interface, where two random images from this set are shown, and the participant is asked to choose the more beautiful of the two. The interface looks similar to Figure

Fig. 5.1 Example of a paiwise comparison, where a user is asked to choose the more beautiful among the two.

5.1, where by one would preferentially end up choosing the image on the right, due to its objectively superior aesthetics as compared to the one on the left. Over the course of time, they gather votes on more than 1.2 pairwise comparisons, distributed across 100k images, given my more than 81k participants. This process, still does not solve the problem of obtaining an annotated dataset of beautiful and ugly looking urban neighbourhoods. For that to happen, we first need to translate these comparisons in to some form of an ordinal ranking of images.

5.2.1 Partitioning the data

To solve the problem of annotating images in terms of beauty, we need to have , at the very least, a sense of relative ranking in terms of most popularly voted to be beautiful to the least. This approach would not quantify beauty in the absolute sense, but in terms of relative consensus, provided we have enough pair wise votes.

To solve the problem of transforming the pairwise voting of the crowd into relative ordinal ranking, we use a popular Bayesian algorithm, used in several multiplayer gaming systems to order leader boards called TrueSkill [HMG07]. This algorithm works by first initializing all the players (which in our case are urban street view images) with equal "skill" (which in our case implies a relative sense of beauty). The algorithm then assumes competitions

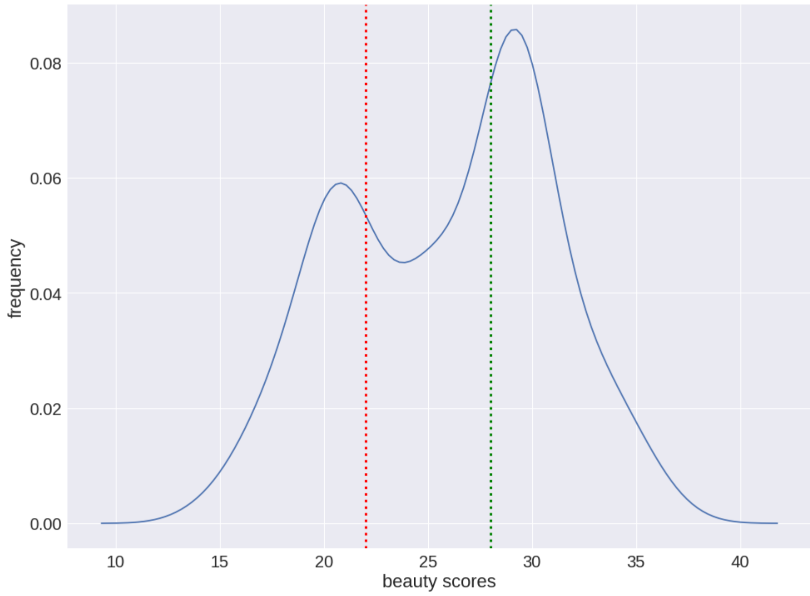


Fig. 5.2 Frequency distribution of beauty scores. The red and green lines represent the thresholds below and above which images are considered ugly and beautiful. Conservatively, images in between are discarded.

among the players in a 1-on-1 fashion. In the case of our work, these competitions are the pairwise votes. Each victory results in addition of some “skill”(beauty) in the victor player’s profile (street-view image) and reduction in “skill” in the loser’s profile. The more skillful the opponent, the higher the rewards in update of one’s skill. With enough average number of competitions played by each player, a steady state order of skilled players emerges. For our dataset, we initialize all images with a “skill” level of 25 and a variance of 3, which signifies the uncertainty in the skill level. This uncertainity would drop, as more competitions are won or lost by any given image. For the sake of accuracy and stability, we filter only the images which have more than 8 votes on them. This reduces our usable dataset from 100k images to just over 20k. But having more than 8 votes, results in a steady state bi-modal distribution of skills as shown in Figure 5.2. This allows us to order the streetview images in an ordinal rank based on relative beauty as seen in Fig 5.3.

Despite having a consensus on the ordering of most of the images, we still end up with some images near the initialized skill level of 25. This means the skills of these images have not been decisively updated through the voting data. To work around these borderline

Fig. 5.3 Sample pairs of street view images ordered by lowest final skill rating on the left to highest on the right.

cases, we partition the distribution of images along two margins in the trueskill space. We conservatively select all images with a Trueskill value less than 22 and assign them to the ugly set of images. We then select all the images with the trueskill value above 28, and assign them to the beauty bin. We arrive at these values of partitioning, by conservatively evaluating how many images we are left with, without polluting the data with borderline cases.

5.2.2 Augmenting the data

Partitioning of the data after evaliating the Trueskill scores on images is a lossy process. By the time we filter images along minimum number of votes, and based on their trueskill scores, we are left with 15,000 street-view images. It has been a well known problem in the area of deep learning, that the algorithm's performance is often limited by the amount of clean curated data available to train on. Another real danger of training on limited data is the phenomenon of over fitting. Overfitting happens when the model under consideration is reasonably complicated with millions of parameters, but the data used to train it is not sufficient to generalize the model's inference. In such cases, the model over fits to the training dataset whereby the model memorises the training data, to reduce the training error. But this model is doomed to perform poorly on a generalized set of data. To avoid this particular set

(a) (b)

Fig. 5.4 Two types of augmentation: (a) rotation of the Street Views camera (based on rotation); and (b) exploration of scenes at increasing distances (based on translation).

of problems, I needed a way to enrich the current high confidence set of steetview images. The enrichment needs to be such that we do not add noise to the dataset, but at the same time increase the diversity of the kind of samples the model is supposed to learn. This means we need more images which are similar but not identical to the two partitions.

The solution I develop involves two approaches. First, we feed each scene's location into the Google Streetview API to obtain the snapshots of the same location at different camera angles[1] (i.e., at $\theta \in -30°, -15°, 15°, 30°$). We assume that any image taken as a variant of a beautiful image at different angles of rotations can safely inherit the label of "beauty" since the field of view only changes by $+-30°$. With this simple (yet conservative) assumption we are able to immediately inflate our dataset by up to a factor of 5. However, the resulting dataset is still too small for robust training. Therefore, again, we feed each scene's location into the Google Streetview API, but we now do so to obtain other scenes at distance $d \in \{10, 20, 40, 60\}$ meters. This will greatly expand our set of scenes, but it might

[1]Google streetview API allows the users to set the bearing and heading of the mast camera, used to acquire the image

do so at the price of introducing scenes whose beauty scores have little to do with the original scene's. This addition of noise may have an adverse effect on the deep learning algorithm's performance for detecting beauty. This addition of translational data, could be done using some heuristics. Imagine a scene I translated in space by 10 meters. The newly acquired scene I_{10}, may be useful in our augmented dataset *iff* the translated scene I_{10} is visually "similar" to the original scene I. The same heuristic metric can be used to either accept or reject any image I_m translated by m meters, into the augmented dataset. For this heuristic test to work, we first need to device a way to compute the "similarity" measure between two images I and I_m. One well known way to evaluate image "similarity" is by represent them in a high dimensional features space, obtained from a pre-trained neural network. This is done by using the activations of a the fully connected layer of a trained convolutional neural network during a forward pass [BL15, LYHC15, VS16]. These activations are then treated as vectors in R^N euclidean space, such that they follow the distance and angle metrics. This allows comparison of images along a similarity metric simply by comparing the cosine distances between their feature vectors. For the sake of similarity of application, we use the best performance version of a pre-trainied PlacesNet deep learning model [ZLX$^+$14]. PlasesNet was trained on streetview images, and is trained with the end goal of classifying streetview images into scene types, such as a beach, highway, garden etc. We represent the two scenes I and I_m with visual features derived from the $FC7$ layer of PlacesNet and compute the similarity between the two corresponding feature vectors using L1 norm. For all scenesI_m at increasing distance $m \in \{10, 20, 40, 60\}$ meters, we take only those whose similarity scores with the original scene is above a threshold. In a conservative fashion, we choose that threshold to be the median similarity between rotated and original scenes (those of the first augmentation step). This assures that the translated images are at the very least as similar to the original, as the rotated images.

Which scenes are more suitable for translation?

To make sure this additional augmentation has not introduced any unwanted noise, we consider two sets of scenes: one containing those that have been taken during this last

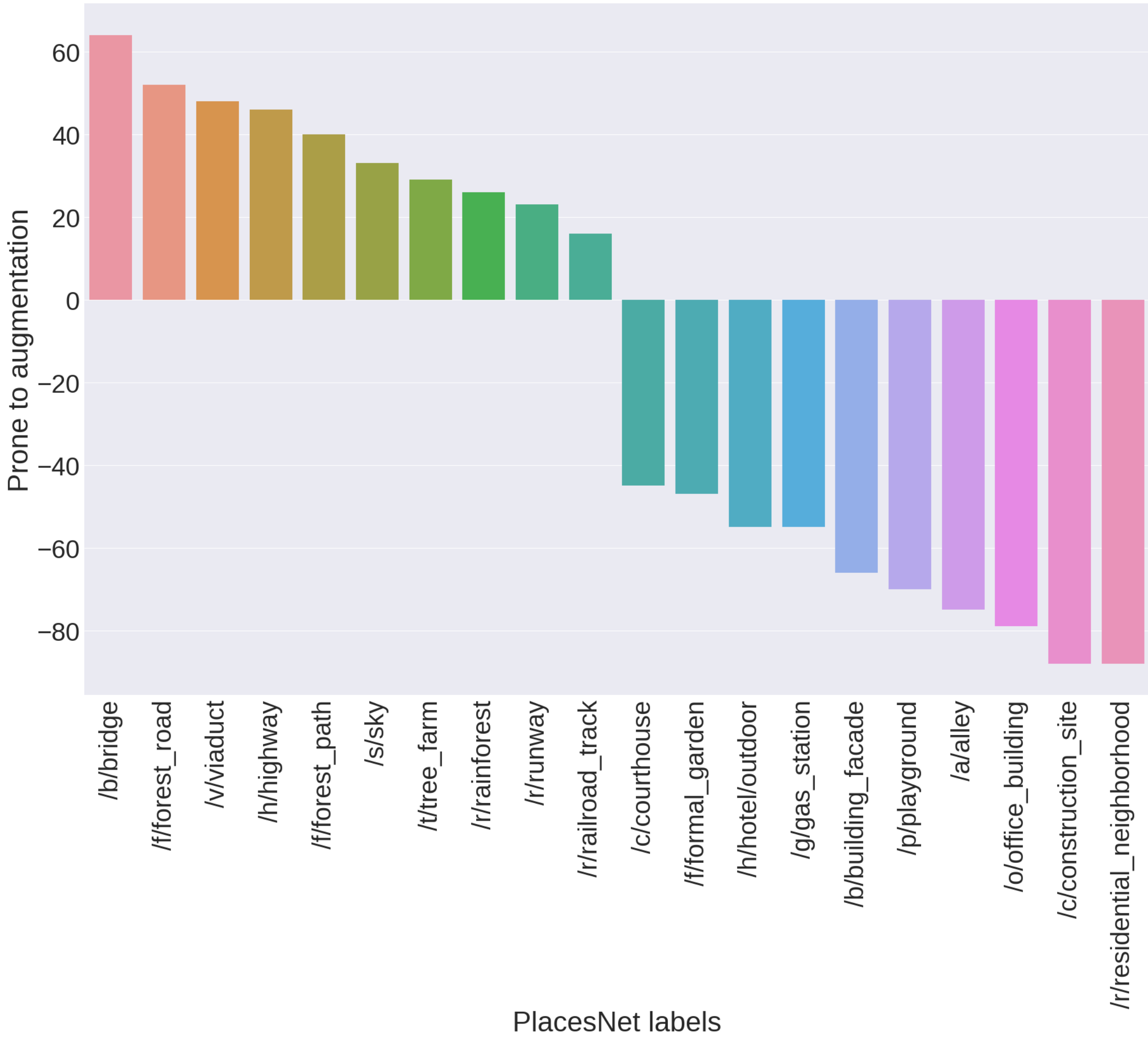


Fig. 5.5 The types of scene that have greater propensity to be correctly augmented with similar scenes at increasing distances.

step, i.e., the one with high similarity to the original scenes:(*taken-set*), and the other containing those that have been filtered away because their similarity metric went above the threshold:(*filtered-set*). Each scene is then scored with PlacesNet [ZLX$^+$14] and is represented with the five most confident scene labels, as per the original output of the model . We then aggregate labels at set level by computing each label's frequency on the *taken-set* and on the *filtered-set*. Finally, we characterize each label's propensity to be correctly augmented as: $\text{prone}(label) = fr(label, taken\text{-}set) - fr(label, filtered\text{-}set)$. This reflects the extent to which a scene with a given scene label is prone to be augmented or not, according to our

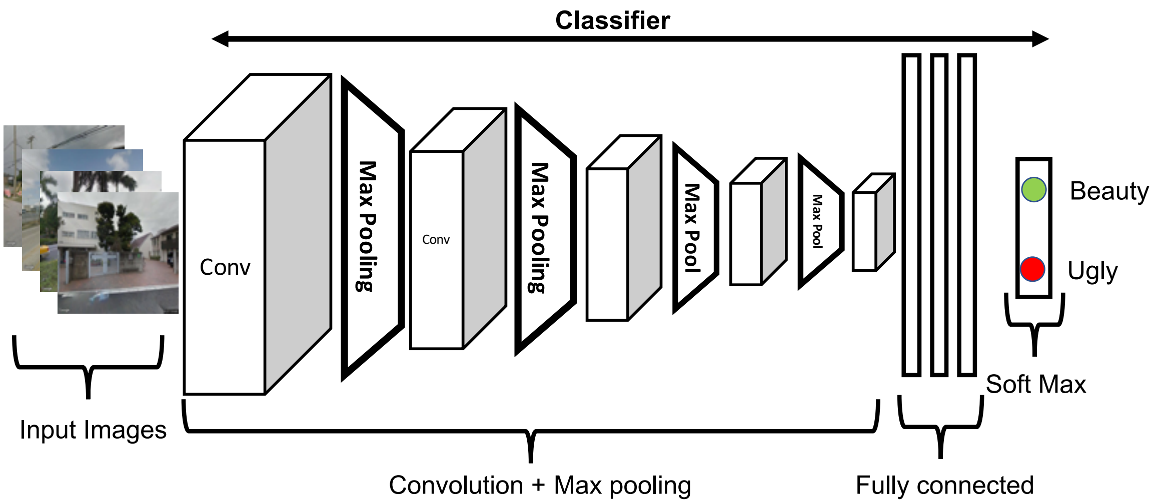


Fig. 5.6

decided threshold method. From Figure 5.5, we find that, as one would expect, scenes that contain high amount of visual continuity over long stretches such as highways, fields and bridges can be augmented at increasing distances while still showing resemblances to the original scene. By contrast, scenes that contain peculiar features with smaller continuity such as gardens, residential neighbourhoods, plazas, and skyscrapers cannot be easily augmented. These scene types also tend to be found in high density parts of the city in which visual diversity within short distances might well be experienced.

5.3 Building a beauty Classifier

Augmentation	Accuracy (Percentage)
None	63
Rotation	68
Rotation + Translation	64
Rotation + Conservative Translation	73.5

Table 5.1 Percentage accuracy for our beauty classifier trained on differently augmented sets of urban scenes.

Having this highly curated set of labeled urban scenes, we are now ready to train classifier C with labels reflecting our beauty assessments. The challenge here is to understand if a deep learning model is able to capture the essence of something as subjective as **perceived urban beauty**.

As for classifier C, we use the CaffeNet architecture, a modified version of AlexNet [K+12, SLJ+15] as seen in Figure 5.6. This has a conventional architecture with 5 convolutional layers; interleaved with max pooling and normalization layers; and terminated by: *(i)* three 4096 dimensional fully connected layers interleaved with dropout layers [SHK+14] (the dropout ratio is set to 0.5 to prevent over-fitting), and *(ii)* by a Softmax layer that classifies the input image into one of two classes of beautiful(1) and ugly(0).

Having C at hand, we now turn to training it. The training is done on a 70% split of the data, and the testing on the remaining 30%. All this is done on increasingly augmented sets of data. We start from our 20k images and progressively augment them with the snapshots obtained with the 5-angle camera rotations, and then with the exploration of scenes at increasing distance $d \in \{10, 20, 40, 60\}$ meters. The idea behind this data augmentation is that the model's accuracy should increase with increasing levels of augmentation. Indeed it does (Table 5.1): it goes from 63% on the set of original scenes to a value as high as 73.5% on the set of fully augmented scenes, which is a notable increase in accuracy for this type of classification tasks. Furthermore, our results match previous ones: for example, Dubey et.al's [DNP+16] model showed an accuracy of 70%, which is comparable to ours. The comparable, and at times better, performance of this classifier with respect to the baseline shows that 1) the notion of subjective beauty can be learn't from a crowd's participation and 2) the novel method of augmenting data, can make deep learning better accessible when the data is sparse and of urban nature.

5.4 Implication of a beauty classifier

In this chapter, we found that using a crowd based pooling of opinions of urban beauty, we can make measurable progress towards quantifying the aesthetic in the context of urban scenes. This too can be done at scale, with a novel way of augmenting sparse datasets. However, just capturing the perception is not enough to make any meaningful contribution towards the urban science. A lot of work has been put in understanding the impact of the urban aesthetic of citizen's health, well being, economical vitality etc. The unanimous consensus points to the fact that cites and their compositions deeply affect our health and well being. At this juncture it is more valuable to first understand if the notion of beauty captured by our deep learning model is analogues to that perceived by real humans. If that is indeed the case, it is worth understanding what aspects of an urban scene are predictive of the perceived beauty. This shall be of immense value to the practitioners of urban design, architecture and urban activism.

In the next chapter, we would explore exactly that. The guiding question is, can the learnt representation of urban beauty be utilized to understand the predictive motifs of urban design for beauty? Can these insights be used in a constructive way to improve existing urban scenes ? And finally, can such a framework be useful to the cause of urban design ?

CHAPTER 6

FROM CROWDS: RECREATING THE URBAN PERCEPTIONS

"What I cannot create, I do not understand" – Richard Feynman

Feynman said this insightful quote in context of understanding deeper concepts in physics. If one cannot teach a concept, one does not understand it deeply. This notion, albeit disconnected with the premise of my thesis, is as relevant as when it comes to building models of any phenomenon. If a model cannot simulate reality, the model has not really captured the essence of what it is trying to learn.

In the last chapter, we saw that a large scale crowd poll about perception of urban spaces, can be used to train a machine learning model that can differentiate between an aesthetically pleasing and unpleasant urban image. We also showed that despite the challenges, we can enrich urban data, using certain spatial heuristics. This however did not address an important question, that is, what is the validity of the learnt representation of beauty? Can we use the model's understanding of beauty to our benefit? Is the model actually learning what humans perceive to be the signatures of beauty? All these questions entail from the final research questions of my dissertation work:

RQ6 *Can we leverage the quantification of the crowd's subjective perception, to improve real spaces?*
RQ7 *Do humans and experts find the improvement realistic?*

We would try to explore the answers to these questions in this chapter, by systematically dissecting the model of beauty learned in the previous chapter. This, as per the definition framed by Feynman, can be done by first using the model to generate what it "thinks" is the notion of urban aesthetic. These generations can be used as recommendations of possible improvements of any given un-aesthetic place. Next we quantify these recommendations using metrics designed around urban design literature. Finally we validated these metrics and recommendations through a user study.

This approach essentially calls back to the DIKW pyramid approach, where by, to derive insights from data, you first create abstractions that come naturally to the structure of data. In this case, the natural abstraction is the structure of the image itself. You are then supposed to represent the knowledge in the abstractions in the form of metrics that could motivate human wisdom. In this case, we believe that the beauty classifier we saw in Chapter 5 has some pertinent knowledge learnt by training on the crowd sourced opinions on beauty. In this chapter, we would develop metrics that operate on this knowledge, to arrive at wisdom inducing insights.

6.1 Related work

Deep learning has been a hot topic and a tool to explore modelling of large datasets in the recent days. The applications of this has seen a steady expansion in the fields other than computer vision and pattern recognition. To list a few, deep learning has made steady headway in natural science allied fields like drug discovery [GHS16], urban science [PZ17, COR$^+$16, LLCG17], health informatics [RWD$^+$16]. The most common forms of deep learning models have been the convolutional neural nets, which we have seen in action in the previous chapter. These convolutional neural networks are essentially layered hierarchical structures of neurons, which can be trained to extract higher dimensional relationships between features in data. This approach has allowed researchers to not only train models to perform simple tasks like detecting objects [K$^+$12] or detecting faces [RPC17], but it has also enabled researchers to quantify much more subjective and intangible properties from images, such as emotions [KLP13],sentiments [B$^+$13] or ambiance of a place [R$^+$15].

Since the introduction of Generative Adversarial Networks (GANs) [GPAM$^+$14], deep learning has been used not only to analyse existing images but also to generate new ones altogether. This family of deep networks has evolved into various forms, from super resolution image generators [LTH$^+$17] to fine-grained in-painting technologies [PKD$^+$16]. Recent approaches have been used to generate images conditioned on specific visual attributes [YYSL15], and these images range from faces [TPW16] to people [MSG$^+$18]. In a similar vein, Nguyen *et al.* [NDY$^+$16] used generative networks to create a natural-looking image that maximizes the activation of a specific neuron in the discriminator. This method was used to bring out the latent representation of an image, that maximizes its probability of a particular class. In theory, the resulting image is the one that “best activates” the neuron under consideration. In practice, it is still a synthetic template that needs further processing to look realistic. Finally, with the recent advancement in Augmented Reality, the application of GANs to generate urban objects in simulated urban scenes have also been successfully shown [AMM$^+$18].

These approaches motivated my work to use GANs, to further investigate the question “What makes an urban scene beautiful?”. To that extent, I design a framework that utilizes GANs in tandem with the beauty classifier, to reconstruct what the classifier presumes to be beautiful. This is done in a way, that provides an incremental change in a target scene to enhance its beauty. Finally, in accordance with the DIWW framework, I develop metrics to better articulate the knowledge learnt by the model. These metrics are novel and are designed around well known metrics in the literature pertaining to urban design and architecture. Finally, I show that this pipeline indeed adds value in the day to day tasks of practitioners of urbanism and the output suggestions around aesthetics of spaces, indeed is perceived to be more aesthetic compared to the input images.

6.2 GAN primer

A new breed of deep learning algorithms have allowed researchers to generate synthetic data samples, based on knowledge and structure learnt form real data samples (x). These family of algorithms are called Generative Adversarial Networks or GANs. GANS work by simultaneously training a Generator $G(z;\theta)$ and a Discriminator $D(x)$ model, who try to out wit each other through training on real samples. The Generator tries to learn to generate the most "real" looking synthetic samples by learning the distribution of x, and the discriminator $D(x)$ tries to learn to maximize the accuracy on guessing which sample comes from a generator and which comes from the real data. At the peak performance, the generator becomes so good, that the discriminator's accuracy cannot go beyond 50% mark, implying a totally random guess. This means the Generator

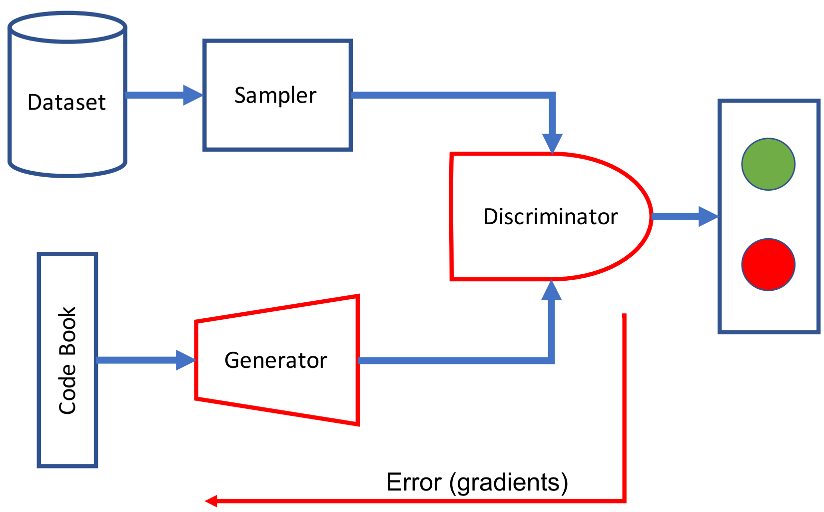


Fig. 6.1 A simplified illustration of the architecture of a Generative Adversarial Network. The red compunets are trained in tandem, one at a time. The discriminator D tries to maximize accuracy to differentiate between I and the generated version $\hat{I}$. And the Generator G tries to minimize accuracy of D to 50% thereby making $\hat{I} \approx I$

The training of this arrangements of neural networks happen in a lock step mode. A simplified version of this arrangement can be found in Figure 6.1. The input to the generator $G(z;\theta)$ is initialized using a prior like a gaussian, which essentially creates a codebook of gaussian noise. More formally the discriminator $D(x)$ tries to maximize assigning a correct probability to each sample x_i to be either from the generator G or the real dataset x. At the same time, the generator G is trying to minimize the losslog$(1-D(G(x)))$. Essentially G and D are locked in a two player min-max game with the value function $V(D,G)$ formalized as :

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\text{ data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))] \tag{6.1}$$

For the premise of my dissertation, I train a generator based on the structure seen at Dosovitsy et.al [DB16], on the entire dataset of streetview images obtained in accordance to the strategies discussed in Chapter 5. The results of the capacity of this generator to approximate streetview scenes can be see from Table 6.1.

6.3 Framework Design

Having the trained classifier at hand and the trained generator of synthetic beautified scenes, the next step is to reconstruct synthetic images using the generator, that maximize the probability of that image being classified as beautiful. The simulated arrangement of this framework can be see in figure 6.2 As a result, given the two classes: ugly y_i and beautiful y_j, the end-to-end model transforms any original scene I_i of class y_i (e.g., ugly scene) into template scene $\hat{I}_j$ that maximizes class y_j (e.g., beautified template scene).

More specifically, given an input image I_i known to be of class y_i (e.g., ugly), our technique outputs $\hat{I}_j$, which is a more beautiful version of it (e.g., I_i is morphed towards the average representation of a beautiful scene) while preserving I_i's details. The technique does so using the "Deep Generator Network for Activation Maximization" (*DGN-AM*) [NDY$^+$16]. Given an input image I_i, *DGN-AM* iteratively re-calculates the color of I_i's pixels in a way the output image $\hat{I}_j$ both maximizes the activation of neuron y_j (e.g., the "beauty neuron") and

Original	Generated
	
	
	

Table 6.1 Examples of our generator's outputs. The original scenes and the generated ones are shown side by side.

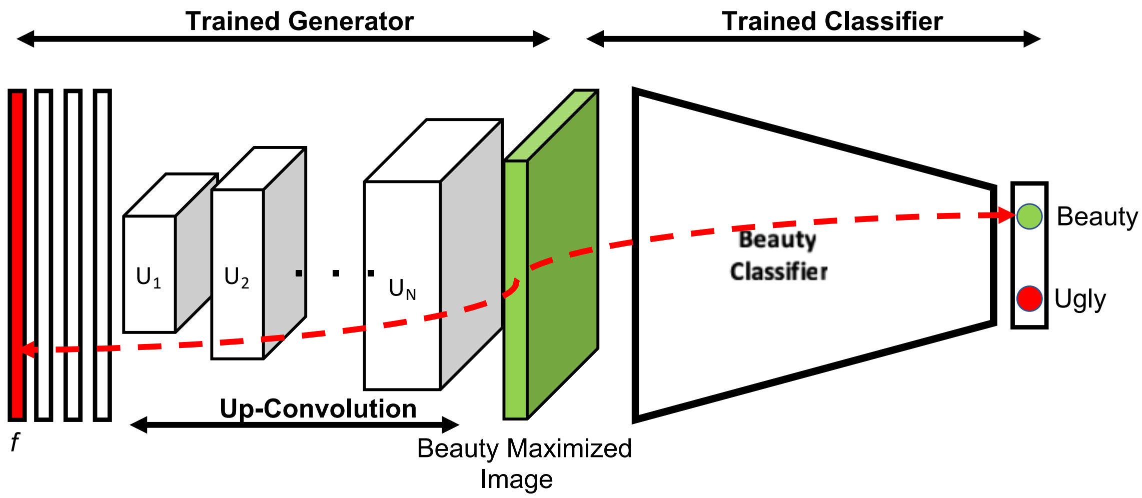


Fig. 6.2 Architecture of the synthetic beauty generator. This consists of a generator of synthetic scenes concatenated with a beauty classifier. The green block is the beauty maximized template $\hat{I}_j$, which is subject to forward and backward passes (red arrow) when optimizing for beauty.

Symbol	**Meaning**
I_i	Original urban scene
Y	Set of annotation classes for urban scenes (e.g., beautiful, ugly)
y_i	Annotation class in Y (e.g., beautiful)
$\hat{I}_j$	Template scene (synthetic image)
I'	Target Image
C	Beauty Classifier

Table 6.2 Notations

looks "photo realistic", which is done by conditioning the maximization to an "image prior". This is equivalent to finding the feature vector f that maximizes the following expression:

$$\hat{I}_j = G(f) : \arg\max_f (C_j(G(f)) - \lambda||f||) \tag{6.2}$$

where:

- $G(f)$ is the image synthetically generated from the candidate feature vector f;
- $C_j(G(f))$ is the activation value of neuron y_j in the scene classifier C (the value to be maximized);

- λ is a L_2 regularization term.

Here the initialization of f is key. If f were to be initialized with random noise, the resulting $G(f)$ would be the average representation of category y_j (of, e.g., beauty). Instead, since f is initialized with the feature vector corresponding to I_i, then the resulting maximized $G(f)$ is I_i's version "morphed to become more beautiful".

The input image is also key. It makes little sense to beautify an already beautiful image, not least because such beautification process would result in a saturated template $\hat{I}_j$ in our framework. For this reason, to generate an image that maximizes the beauty neuron in the classifier C, we restrict the corresponding input image to be in class y_i (i.e., ugly scenes as per the divisions in Figure 5.2). We do the opposite when maximizing the ugly neuron.

We now have template scene $\hat{I}_j$ (which is a synthetic beautified version of original scene I_i) and need to retrieve a realistic looking version of it. We do so by: *i)* representing each of our original scenes in Step 1 (including $\hat{I}_j$) as a 4096 dimensional feature vector derived from the FC7 layer of the PlacesNet [ZLX$^+$14]; *ii)* computing the distance (as L_2 Norm) between $\hat{I}_j$'s feature vector and each of the original scene's feature vector; and *iii)* selecting the original scene most similar (smaller distance) to $\hat{I}_j$. This results into the selection of the beautified scene I_j.

6.4 Motifs of urban beauty

Since original scene I_i and beautified scene I_j are real scenes with the same structural characteristics (e.g., point of view, layout), we can easily compare them in terms of presence or absence of urban elements extracted by computer vision tools such as SegNet [BKC15] and PlacesNet [ZLX$^+$14]. That is, we can determine how the original scene and its beautified version differ in terms of urban design elements.

Original (I_i)	Latent Beauty representation ($\hat{I}_j$)	Beautified (I_j)
		
		
		
		
		

Table 6.3 Examples of the “FaceLifting” process, which tends to add greenery, narrow roads, and pavements.

6.5 Evaluation

The goal of FaceLift is to transform existing urban scenes into versions that: *i)* people perceive more beautiful; *ii)* contain urban elements typical of great urban spaces; *iii)* are easy to interpret; and *iv)* architects and urban planners find useful. To ascertain whether FaceLift meets that composite goal, we answer the following questions next:

Q1 Do individuals perceive "FaceLifted" scenes to be beautiful?

Q2 Does our framework produce scenes that possess urban elements typical of great spaces?

Q3 Which urban elements are mostly associated with beautiful scenes?

Q4 Do architects and urban planners find FaceLift's insights useful?

6.5.1 Q1 People's perceptions of beautified scenes

To ascertain whether "FaceLifted" scenes are perceived by individuals as they are supposed to, we run a crowd-sourcing experiment on Amazon Mechanical Turk. We randomly select 200 scenes, 100 beautiful and 100 ugly (taken at the bottom 10 and top 10 percentiles of the Trueskill's score distribution of Figure 5.2). Our framework then transforms each ugly scene into its beautified version, and each beautiful scene into its corresponding 'uglified'. These scenes are arranged into pairs, each of which contains the original scene and its beautified or uglified version. On Mechanical Turk, we only select verified masters as our crowd-sourcing workers (those with an approval rate above 90% during the past 30 days), pay them $0.1 per task, and ask each of them to choose the most beautiful scene for each given pair. We make sure to have at least 3 votes for each scene pair. Overall, our workers end up selecting the scenes that are actually beautiful 77.5% of the times, suggesting that "FaceLifted" scenes are indeed perceived to be more beautiful by people.

6.5.2 Q2 Are beautified scenes great urban spaces?

To answer that question, we need to understand what makes a space great. After reviewing the literature in urban planning, we identify four factors associated with great places [EC13, Ale77] (Table 6.4): they mainly tend to be walkable, offer greenery, feel cozy, and be visually rich.

Metric	Description
Walkability	Walkable streets support people's natural tendency to explore spaces [EC13, QASD15, Spe12].
Green Spaces	The presence of greenery has repeatedly been found to impact people's well-being [Ale77]. Under certain conditions, it could also promote social interactions [QOC14]. Not all types of greenery have to be considered the same though: dense forests or unkempt greens might well have a negative impact [Jac61].
Landmarks	Feeling lost is not a pleasant experience, and the presence of landmarks have been shown to contribute to the legibility and navigability of spaces [Lyn60, QOC14, EC13, QPAC13].
Privacy-Openness	The sense of privacy conveyed by a place's structure (as opposed to a sense of openness) impacts its perception [EC13].
Visual Complexity	Visual complexity is a measure of how diverse an urban scene is in terms of design materials, textures, and objects [EC13]. The relationship between complexity and preferences generally follows an 'inverted-U' shape: we prefer places of medium complexity rather than places of low or high complexity [Ulr83].

Table 6.4 Urban Design Metrics

To automatically extract visual cues related to these four factors, we select 500 ugly scenes and 500 beautiful ones at random, transform them into their opposite aesthetic qualities (i.e., ugly ones are beautified, and beautiful ones are 'uglified'), and compare which urban elements related to the four factors distinguish uglified scenes from beautified ones.

We extract labels from each of our 1,000 scenes using two image classifiers. First, using PlacesNet [ZLX$^+$14], we label each of our scenes according to a classification containing

205 labels (reflecting, for example, landmarks, natural elements), and retain the five labels with highest confidence scores for the scene. We then manually classify these 205 labels into the 4 built environment types namely *Walkable, Architectural, Natural and Landmark*. These types are inspired by the guidelines for measuring urban design [EC13]. The specifics of their category definitions and the actual labels can be seen in Appendix 2(Chapter B). Second, using Segnet [BKC15], we label each of our scenes according to a classification containing 12 labels. Segnet is trained on dash-cam images, and classifies each scene pixel with one of these twelve labels: road, sky, trees, buildings, poles, signage, pedestrians, vehicles, bicycles, pavement, fences, and road markings.

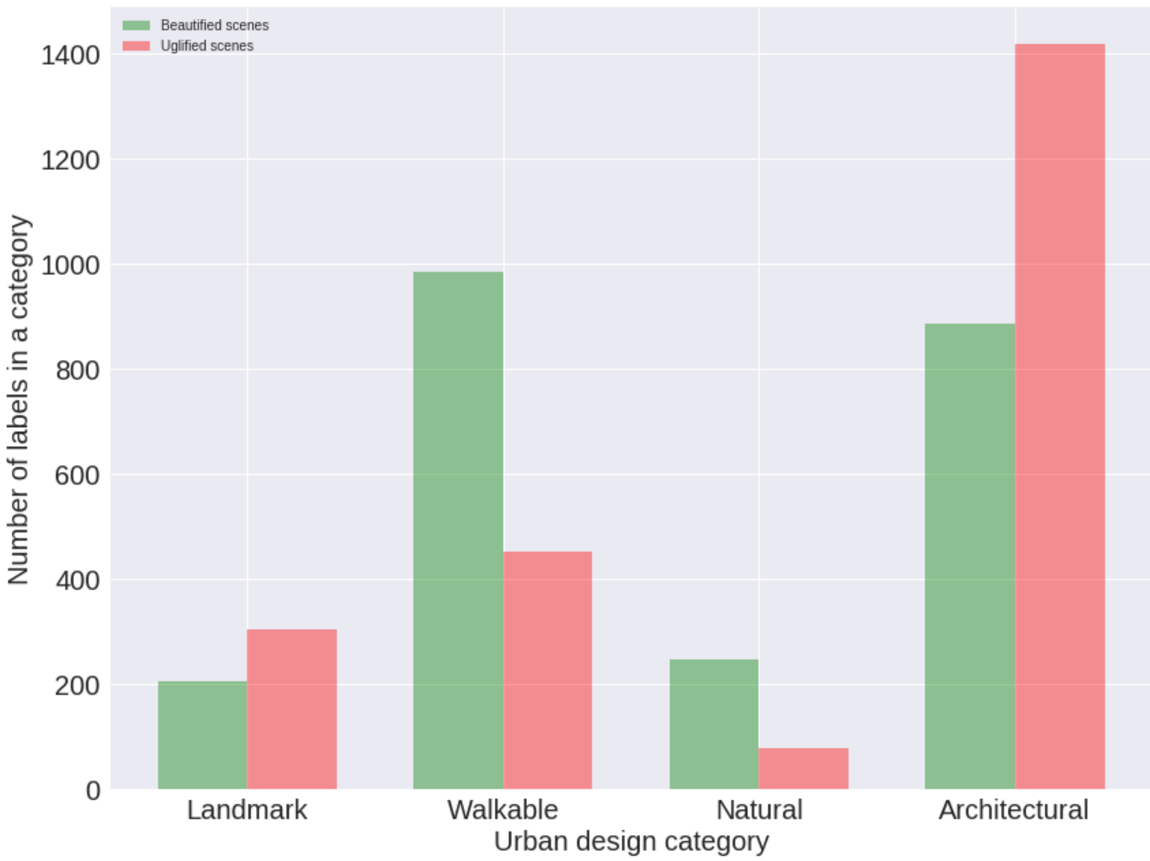


Fig. 6.3 Number of labels in specific urban design categories (on the *x*-axis) found in beautified scenes as opposed to those found in uglified scenes.

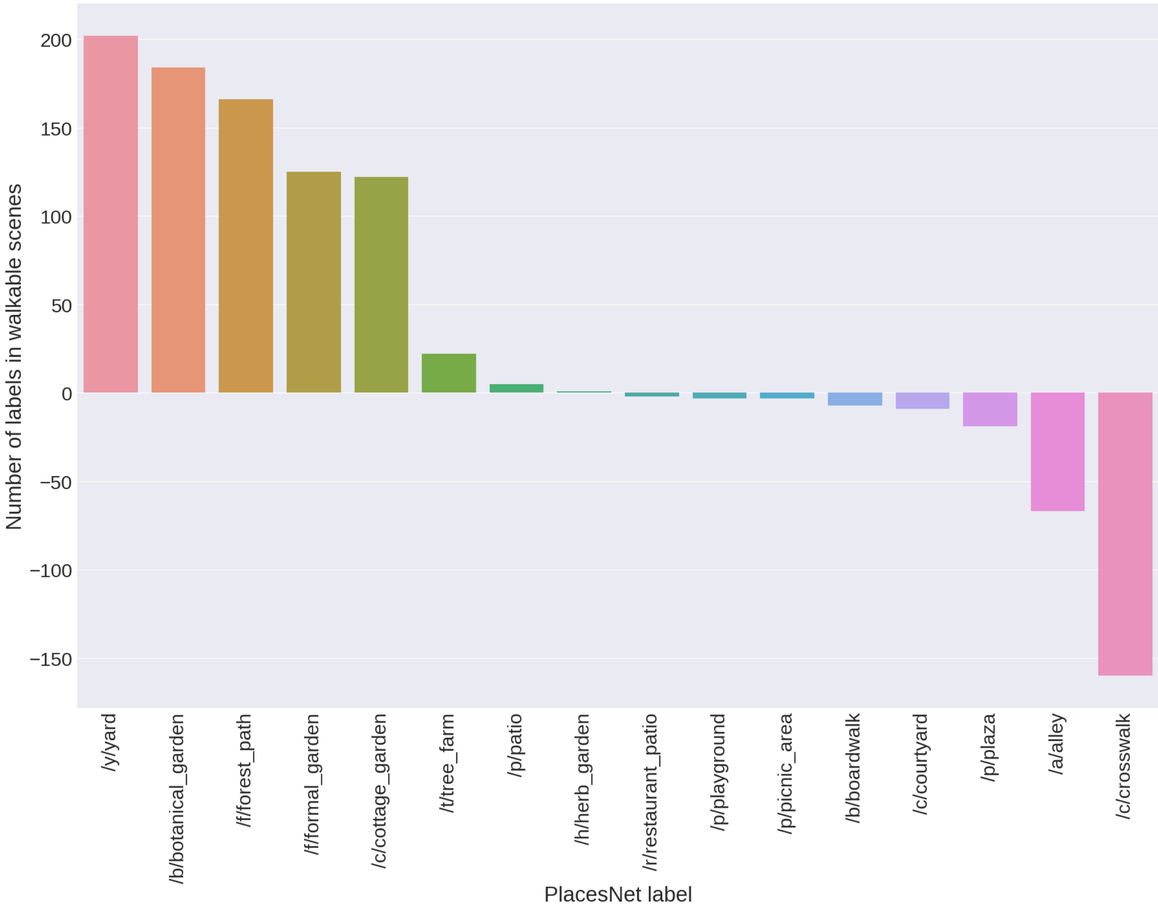


Fig. 6.4 Count of specific walkability-related labels (on the x-axis) found in beautified scenes minus the count of the same labels found in uglified scenes.

Having these two ways of labelling scenes, we can now test whether the expectations set by the literature describing metrics of great urban spaces (Table 6.4) are met in the Face-lifted scenes.

H1 Beautified scenes tend to be walkable. We manually select only the PlacesNet labels that are related to walkability. These labels include, for example, *abbey, plaza, courtyard, garden, picnic area, and park*. To test hypothesis *H1*, we count the number of walkability-related labels found in beautified scenes as opposed to those found in uglified scenes (Figure 6.3): the former contain twice as many walkability labels than the latter. We then determine which types of scenes are associated with beauty (Figure 6.4). Unsurprisingly, beautified scenes

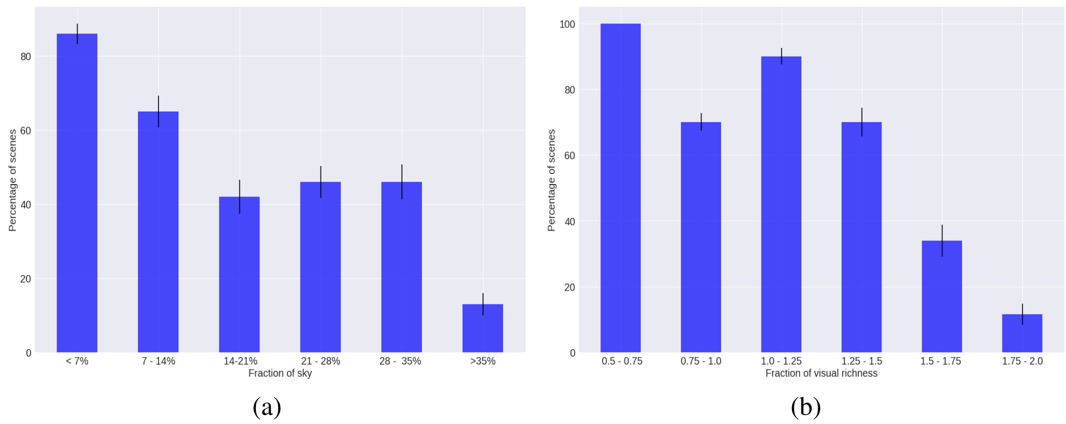


Fig. 6.5 The percentage of beautified scenes (y-axis): (a) having an increasing presence of sky (on the x-axis); and (b) having an increasing level of visual richness (on the x-axis). The error bars represent standard errors obtained by random re-sampling of the data for 500 iterations.

tend to show gardens, yards, and small paths. By contrast, uglified ones tend to show built environment features such as shop fronts and broad roads.

H2 Beautified scenes tend to offer green spaces. We manually select only the PlacesNet labels that are related to greenery. These labels include, for example, *fields, pasture, forest, ocean, and beach*. Then, in our 1,000 scenes, to test hypothesis *H2*, we count the number of nature-related labels found in beautified scenes as opposed to those found in uglified scenes (Figure 6.3): the former contain more than twice as many nature-related labels than the latter. To test this hypothesis further, we compute the fraction of 'tree' pixels (using SegNet's label 'tree') in beautified and uglified scenes, and find that beautification adds 32% of tree pixels, while uglification removes 17% of them.

H3 Beautified scenes tend to feel private and 'cozy'. To test hypothesis *H3*, we count the fraction of pixels that Segnet labeled as 'sky' and show the results in a bin plot in Figure 6.5a: the x-axis has six bins (each of which represents a given range of sky fraction), and the y-axis shows the percentage of beautified *vs.* uglified scenes that fall into each bin. Beautified scenes tend to be cozier (lower sky presence) than the corresponding original scenes.

Pair of urban elements	β_1	β_2	β_3	Error Rate (Percentage)
Buildings - Trees	-0.032	0.084	0.005	12.7
Sky - Buildings	-0.08	-0.11	0.064	14.4
Roads - Vehicles	-0.015	-0.05	0.023	40.6
Sky - Trees	0.03	0.11	-0.012	12.8
Roads - Trees	0.04	0.10	-0.031	13.5
Roads - Buildings	-0.05	-0.097	0.04	20.2

Table 6.5 Coefficients of logistic regressions run on one pair of predictors at the time.

H4 Beautified scenes tend to be visually rich. To quantify to which extent scenes are visually rich, we measure their visual complexity [EC13] as the amount of disorder in terms of distribution of (Segnet) urban elements in the scene:

$$H(X) = -\sum p(i) \log p(i) \tag{6.3}$$

where i is the i^{th} Segnet's label. The total number of labels is twelve. The higher $H(X)$, the higher the scene's entropy, that is, the higher the scene's complexity. It has been suggested that the relationship between complexity and pleasantness follows an 'inverted U' shape [Ulr83]: we prefer places of medium complexity rather than places of low or high complexity. To test that, we show the percentage of beautified scenes that fall into each complexity bin (Figure 6.5b): we do not find a strong evidence of the 'inverted U' shape hypothesis, in that, beautified scenes are of low to medium complexity, while uglified ones are of high complexity.

6.5.3 Q3 Urban elements of beautified scenes

To determine which urban elements are the best predictors of urban beauty and the extent to which they are so, we run a logistic regression, and, to ease interpretation, we do so on one pair of predictors at the time:

$$Pr(\text{beautiful}) = logit^{-1}(\alpha + \beta_1 * V_1 + \beta_2 * V_2 + \beta_3 * V_1.V_2) \tag{6.4}$$

Use case	Definitely Not	Probably Not	Probably	Very Probably	Definitely
Decision Making	4.8%	9.5%	38%	28.6%	19%
Participatory Urban Planning	0%	4.8%	52.4%	23.8%	19%
Promote Green Cities	4.8%	0%	47.6%	19%	28.6%

Table 6.6 Urban experts polled about the extent to which an interactive map of "FaceLifted" scenes promotes: (a) decision making; (b) citizen participation in urban planning; and (c) promotion of green cities

where $V1$ is the fraction of the scene's pixels marked with one Segnet's label, say, "buildings" (over the total number of pixels), and $V2$ is the fraction of the scene's pixels marked with another label, say, "trees". The result consists of three beta coefficients: β_1 reflects $V1$'s contribution in predicting beauty, β_2 reflects $V2$'s contribution, and β_3 is the interaction effect, that is, it reflects the contribution of the dependency between $V1$ and $V2$ in predicting beauty. We run logistic regressions on the five factors that have been found to be most predictive of urban beauty [QOC14, EC13, Ale77], and show the results in Table 6.5.

Since we are using logistic regressions, the quantitative interpretation of the beta coefficients is eased by the "divide by 4 rule" [Vau08]: we can take the β coefficients and "divide them by 4 to get an upper bound of the predictive difference corresponding to a unit difference" in beauty [Vau08]. For example, take the results in the first row of Table 6.5. In the model $Pr(beautiful) = logit^{-1}(\alpha - 0.032 \cdot buildings + 0.084 \cdot trees + 0.005 \cdot buildings \cdot trees)$, we can divide - 0.032/4 to get -0.008: a difference of 1 in the fraction of pixels being buildings corresponds to no more than a 0.8% *negative* difference in the probability of the scene being beautiful. In a similar way, a difference of 1 in the fraction of pixels being trees corresponds to no more than a 0.021% *positive* difference in the probability of the scene being beautiful. By considering the remaining results in Table 6.5, we find that, across all pairwise comparisons, trees is the most positive element associated with beauty, while roads and buildings are the most negative ones. These results match previous literature in urban design of what makes spaces great, adding further external validity to our framework's beautification.

6.5.4 Q4 Do architects and urban planners find it useful?

To ascertain whether practitioners find FaceLift potentially useful, we build an interactive map of the city of Boston in which, for selected points, we show pairs of urban scenes

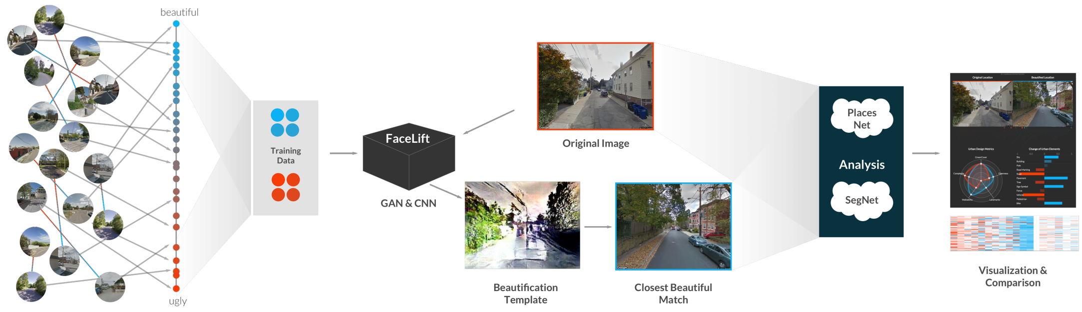


Fig. 6.6 An end to end illustration of the FaceLift framework.

before/after beautification (Figure 6.7)[1]. We then send that map along with a survey to 20 experts in architecture, urban planning, and data visualization around the world. Questions were asked with a non-neutral response Likert scale (Table 6.6). That is because previous work [BFT12, Moo08] has shown that such a scale: *(i)* pushes respondents to "take a stance", given the absence of a neutral response; and *(ii)* works best if respondents are experts in the subject matter of the survey as responses of the "I don't know" type tend to be rare (as it is has been the case for our survey). The experts had to complete tasks in which they rated FaceLift based on how well it supports decision making, participatory urbanism, and the promotion of green spaces. According to our experts (Table 6.6), the tool can very probably supports decision making, probably support participatory urbanism, and definitely promote green spaces. These results are also qualitatively supported by our experts' comments, which include: "*The maps reveal patterns that might not otherwise be apparent*", "*The tool helps focusing on parameters to identify beauty in the city while exploring it*", and "*The metrics are nice. It made me think more about beautiful places needing a combination of criteria, rather than a high score on one or two dimensions. It made me realize that these criteria are probably spatially correlated*".

[1]It is worth noting that the visualization is not claimed to be my original contribution. The design of the interactive map was done by the first author of article 5

Fig. 6.7 Interactive visualization of FaceLifted scenes in Boston

6.6 Conclusion

FaceLift is a framework that automatically beautifies urban scenes by combining recent approaches of Generative Adversarial Networks and Deep Convolutional Networks. To make it usable by practitioners, the framework is also able to explain which urban elements have been added/removed during the beautification process.

There are still important limitations though. One is data bias. The framework is as good as its training data, and more work has to go into collecting reliable ground truth data on human perceptions. This data should ideally be stratified according to the people's characteristics that impact their perceptions. The other main limitation is that generative models are hard to control, and more work has to go into offering principled ways of fine-tuning the generative process.

Despite these limitations, FaceLift has the potential to support urban interventions in scalable and replicable ways: it can be applied to an entire city (scalable), across a variety of cities (replicable). To turn existing spaces into something more beautiful, that will still be the duty of architecture. Yet, with technologies similar to FaceLift more readily available, the complex job of recreating restorative spaces in an increasingly urbanized world will be greatly simplified.

After all, "we delight in complexity to which genius have lent an appearance of simplicity." [DB08] In the context of future work, that genius is represented by future technologies that will help us deal with the complexity of our cities.

6.6.1 Limitations and biases

Like any supervised deep learning based framework, this work is only able to learn what is present in the data. Hence the method of acquiring annotations for urban images can introduce huge biases in the model. The current model is trained on images acquired from the study on streetscore [NPRH14]. However their annotation is open to general public and there is not way we can remove biases that come with culture and location, in a highly subjective effect like beauty. Moreover because the pair wise choice is simply done by clicking one of the two images, the data might have noise introduced by non-serious participants. Such biases are bound to be picked up by the deep learning model. One can argue that the preference of our model for greenery , is a form of bias in the data. Another bias introduced because of data is the model's lack of preference to pedestrians. This bias was established well in advance because Google tries to remove most of the people from their street view images for privacy reasons. Hence people, which make up a major aspect of urban vitality, are completely missing from most dataset images and hence from the facelift transformations. Another Limitation of our work is in the metric formation. The computational metrics developed to capture the real urban design metrics are designed using heuristics. There needs to be more crowd and expert validation to establish the validity of their formulation.

6.6.2 Implications

In the works through my Ph.D., the idea of quantifying the subjective has always been the central theme. The subjective is where most of our human existence takes place. As Emanuel Kant instructs, “Look closely, the beautiful may be small”. And indeed, the insights that drive the quantification of beauty in this part of my work, seem to be almost intuitive. Greenery, walking spaces and open spaces seem drive our perception of the aesthetic, which may be grounded in our natural roots and evolutionary need to be safe and close to nature. Despite of the root causes seeming intuitive, the implications of these simple properties of urban spaces on mental and physical health seem to be far more crucial in today’s world. At such a juncture, it is worth while to develop these frameworks to capture human subjective experiences. Thankfully because of the convergence in our online and offline lives, we are generating far more data about our offline experiences, online. Data acquired from how we interact with different on-line services, may very well improve our off-line lives.

CHAPTER 7

CLOSING NOTES AND OUTLOOK FOR THE FUTURE

> "The unity of all science consists alone in its method, not in its material ... It is not the facts themselves which form science, but the method in which they are dealt with." - Karl Pearson

The guiding principle for this dissertation has been understanding how human subjective opinions can be leveraged for social good. The aim was to advance the understanding of how something as subjective as expression of support or perception of beauty, can be quantified. I explored two realms of this problem, one dealing with groups of humans forming communities around a supportive cause, and other looking at a vast set of un-related people expressing their subjective opinions about urban aesthetics. In both cases, the key was the set of methods used to tease out the signatures of these subjective qualities from data. To that end, Karl Pearson's quote is very apt and encapsulates the key contributions of my work.

Working with a framework, of first acquiring and curating **Data**, then building key abstractions on top to capture the **Information**, then building metrics to extract the **Knowledge** allowed me to build a pipeline to work with data from a diverse set of applications. The hope is that this pipeline can be used to generate **Wisdom** that drives interventions

The work done in my dissertation is going to inspire the vision for my research for the next few years, and I would like to close this journey by walking through a few aspects of what I feel is the path forward.

7.1 Open problems

Throughout this dissertation, I came across interesting problems which I would have loved to investigate, but couldn't in the interest of focus and time. I would like to enumerate a few as an exercise in reflection. At the very least, I would like to discuss the problems and illuminate them for further investigation by the community.

7.1.1 Triadic closures in conversation graphs

Triadic closure has been shown to be an important mechanism in the literature of social networks [Gra77, MVF11] through which social ties get established. The mechanism has also been widely explored as a measure to for recommendation systems [ST14, LTH$^+$13].

Despite this wide prevalence in social networks, dialogue structures on reddit in the context of social support seem to exhibit a different behaviour , as seen in Chapter 4. As seen from the results in Appendix A, motifs that resemble a triadic closure are very rare in either cases. However the precursors to a triadic closure, such as 201 variants, 111 variants and 021 variants are all expressed with statistical significance in the baseline as well as supportive conversations. Triadic closures have shown to be vital for information diffusion in networks [BGV$^+$16]. Hence investigating closures in dialogue structures on reddit could lead to some interesting counter intuitive mechanisms.

7.1.2 Colouring ties in conversation graphs

In my dissertation I explored the utility of language in capturing the strength of ties in a dialogue graph (as seen in Chapter 4). However, it is important to note that just measuring alignment between two exchanges might be resulting a loss of crucial information. Linguistically, it would be worthwhile to capture the essence of dialogues along a multi-dimensional scale, since that is how actual exchanges take place. To that end, it would be of value to colour the links using affective components of an exchange between two people, such as empathy, affection, trust, anxiety, animosity, friendship etc. These colours could further help us develop methods to detect toxic behaviours online and capture the overall tone of a discussion.

7.1.3 Exploring cultural biases in subjective perception

An important limitation in using the crowd's opinion to quantify the subjective is the trade-off between data volume and data bias. To train any reliable deep learning model, you need a reasonable volume of data to train. This limits the amount of stratification one can do in the crowd opinions along cultural, geographical and social lines. Stratifying data further could lead to over-fitted models. But using the bulk of data as one monolithic chuck endangers the model to learn the least common denominator in the subjective preferences. That means the cultural nuances about the concept of beauty that enrich our world are all averaged out. Indeed as seen in Chapter 6 the FaceLift model tends to associate foliage with beauty to a high degree. There are methods in the literature to solve these problems, by teasing out biases in the models by partitioning the data in a clever manner.

7.2 Future Outlook

In this Ph.D I aimed at exploring and exploiting the potential of machine learning, to understand how the subjective can be sensed from web scale data. To that end, it is worth discussing about how the methods I developed align with the research I intend to pursue.

In the future I would like to extend this work along two key dimensions which are based on a shared theme *"Capitalizing on the subjective perceptions to deliver impactful interventions"*.

7.2.1 Empathic healthcare

Improving healthcare to provide an empathic experience to patients in an economical and scalable way is the most crucial challenge of 21st century. The world saw an explosive growth in population during the baby boomer generation. This same cohort has now become the largest ageing cohort in the history of the world. On account of this and the rise of chronic diseases, the health care sector has seen unprecedented growth in the past decade[1]. With growth, comes the challenges of scale. Despite ever increasing investment[2], the UK's NHS still is riling under the pressure of rising patient numbers, dwindling staff and longer wait times [May18]. Longer wait times also imply that doctors are on an average spending less time with the patients. This has taken a serious toll on the doctor patient empathic communication. It has been shown that the doctor patient relationship plays a vital role in accuracy of diagnosis of the disease ,prognosis of the patient and overall satisfaction of the patient [JBS⁺11, Ben91].

At such a juncture, there is a rising need to solve frictions along these points of contacts for the NHS. At the same time, it is extremely important to provide psyco-social infrastructure for this ageing population, especially at times when the ailments are chronic in nature, and the social support structures are fragile. The vision for an empathic healthcare, puts forth a framework where the psycho-social aspects of health are put at the centre of the system, along with the biological well being. This can be done by re-engineering several pipelines, through which a patient engages with the healthcare systems. This means that a treatment plan is not just a relationship between the doctor and the patient, but involves a support structure of both AI-driven agents, peers, online-volunteers and healthcare workers. My work done with the Chronic Obstructive Pulmonary Disease community shows that patients of chronic diseases tend to thrive as a part of an online social support community, and at times can take up the mantle to provide crucial information [JSC⁺18]. My work with the suicide watch support community also shows that there are quantifiable structures of support which exhibit a patient centric structure of engagement. I foresee that with the help of A.I. models trained on these empathic interactions, we could one day see virtual support groups which are supported by the healthcare provider and allow self and group management of chronic conditions. This would allow the healthcare providers to save costs in terms of lost appointment times for patients of chronic conditions. But at the same time, it would provide

[1] https://www2.deloitte.com/global/en/pages/life-sciences-and-healthcare/articles/global-health-care-sector-outlook.html

[2] https://www.bbc.co.uk/news/health-46524257

a true bio-psycho-social framework for chronic symptom management. Overall, I see that empathic signatures in communications, can be learnt at scale, provided that we can pinpoint where exactly such interactions are happening. And I think my work paves a path forward for that to happen.

7.2.2 Perceptive urbanism

Urbanism is a term used commonly to describe works that deal with problems and mechanisms that shape an urban environment. There has been a sharp rise in research in the field of urbanism due to the age of open data. Many of these works look at open urban data about socio-economic indicators in conjunction with social outcome variables like health [SBL$^+$12, VQC$^+$15], well being [CSH$^+$17], quality of air [MD03] etc. Some more inter-disciplinary works have also shown that built environment in cities resonate with our personality and perception of qualities like safety, richness and beauty [DNVZ$^+$16b, DNP$^+$16].Some other studies have also shown that something as simple as presence of green canopy can be correlated with reduced depression related prescriptions [HKR$^+$18].

In the same vein, my work done on crowd based urban design, allowed me to leverage perception of the crowds to improve urban spaces. I believe the most natural extension in this case is utilizing these methods to explore further linkages between our urban environment and our quantified self. Questions like "how does the city influence our mental health?" or "how does built environment influence the air we breath?" are now within the reach of data science. The possibility of linking perception driven metrics in urban spaces, with socio-economic or health indicators is what I define as perceptive urbanism. This aspect of urbanism could actually be immensely helpful in developing interventions with minimum costs but maximum impact.

7.3 Conclusion

In the hindsight, this work has been a result of a series of fortunate accidents which allowed me to explore and exploit interesting topics in the fields of information retrieval, network science and at times social sciences. Being a computer scientist in today's age, I believe, is as much an exercise in a broad understanding of today's social problems, as it is about technical details and methods to solve them. My Ph.D. has given me an unfettered opportunity to spend my time and resources in broadening my horizons. I hope I keep the pursuit alive in the due course of my career.

REFERENCES

[AERA14] Sean M Arietta, Alexei A Efros, Ravi Ramamoorthi, and Maneesh Agrawala. City forensics: Using visual elements to predict non-visual city attributes. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2624–2633, 2014.

[AFGG15] Kathina Ali, Louise Farrer, Amelia Gulliver, and Kathleen M Griffiths. Online peer-to-peer support for young people with mental health problems: a systematic review. *JMIR mental health*, 2(2):e19, 2015.

[AH00] Arthur Armstrong and John Hagel. The real value of online communities. *Knowledge and communities*, 74(3):85–95, 2000.

[AJB00] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *nature*, 406(6794):378, 2000.

[AJS+13] Mohammad Adibuzzaman, Niharika Jain, Nicholas Steinhafel, Munir Haque, Ferdaus Ahmed, Sheikh Ahamed, and Richard Love. In situ affect detection in mobile devices: a multimodal approach for advertisement using social network. *ACM SIGAPP Applied Computing Review*, 13(4):67–77, 2013.

[Ale77] Christopher Alexander. *A Pattern Language: Towns, Buildings, Constructions*. Oxford University Press, 1977.

[AMM+18] Hassan Abu Alhaija, Siva Karthik Mustikovela, Lars Mescheder, Andreas Geiger, and Carsten Rother. Augmented reality meets computer vision: Efficient data generation for urban driving scenes. *International Journal of Computer Vision*, 126(9):961–972, 2018.

[AVKR16] Chris Allen, Ivaylo Vassilev, Anne Kennedy, and Anne Rogers. Long-term condition self-management support in online communities: a meta-synthesis of qualitative papers. *Journal of medical Internet research*, 18(3), 2016.

[B+13] Damian Borth et al. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *Proceedings of the 21st ACM MM*, MM '13, New York, NY, USA, 2013. ACM.

[BAH+86] George W Brown, Bernice Andrews, Tirril Harris, Zsuzsanna Adler, and L Bridge. Social support, self-esteem and depression. *Psychological medicine*, 16(4):813–831, 1986.

[BBLO01] Kylie Ball, Adrian Bauman, Eva Leslie, and Neville Owen. Perceived environmental aesthetics and convenience and company are associated with walking for exercise among australian adults. *Journal of Preventive Medicine*, 33(5):434–440, 2001.

[BC11] Yvonne Kiera Bartlett and Neil S Coulson. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. *Patient education and counseling*, 83(1):113–119, 2011.

[BCH+90] Robert S Baron, Carolyn E Cutrona, Daniel Hicklin, Daniel W Russell, and David M Lubaroff. Social support and immune function among spouses of cancer patients. *Journal of Personality and Social Psychology*, 59(2):344, 1990.

[BDIF14] Ginestra Bianconi, Richard K Darst, Jacopo Iacovacci, and Santo Fortunato. Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E*, 90(4):042806, 2014.

[BDSZ16] Alfredo Braunstein, Luca Dall'Asta, Guilhem Semerjian, and Lenka Zdeborová. Network dismantling. *Proceedings of the National Academy of Sciences*, 113(44):12368–12373, 2016.

[Ben91] Jozien Bensing. Doctor-patient communication and the quality of care. *Social science & medicine*, 32(11):1301–1310, 1991.

[BFT12] Aphrodite Baka, Lia Figgou, and Vasiliki Triga. 'neither agree, nor disagree': a critical analysis of the middle answer category in voting advice applications. *International Journal of Electronic Governance*, 5(3-4):244–263, 2012.

[BGV+16] Mahmoudreza Babaei, Przemyslaw Grabowicz, Isabel Valera, Krishna P Gummadi, and Manuel Gomez-Rodriguez. On the efficiency of the information networks in social media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 83–92. ACM, 2016.

[BKC15] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*, 2015.

[BL15] Artem Babenko and Victor Lempitsky. Aggregating local deep features for image retrieval. In *Proceedings of the IEEE international conference on computer vision*, pages 1269–1277, 2015.

[BLF01] Tim Berners-Lee and Mark Fischetti. *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*. DIANE Publishing Company, 2001.

[BM01] Vladimir Batagelj and Andrej Mrvar. A subquadratic triad census algorithm for large sparse networks with small maximum degree. *Social Networks*, 23(3):237–243, 2001.

[BMBL09] Stephen P Borgatti, Ajay Mehra, Daniel J Brass, and Giuseppe Labianca. Network analysis in the social sciences. *science*, 323(5916):892–895, 2009.

[Bor04] Maria Borges. What can kant teach us about emotions? *The Journal of Philosophy*, 101(3):140–158, 2004.

[BSBMDS16] Chantal Balasooriya-Smeekens, Andrew Bateman, Jonathan Mant, and Anna De Simoni. Barriers and facilitators to staying in work after stroke: insight from an online forum. *BMJ open*, 6(4):e009974, 2016.

[BSG14] Saeideh Bakhshi, David A Shamma, and Eric Gilbert. Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 965–974. ACM, 2014.

[Bur04] Ronald S Burt. Structural holes and good ideas. *American journal of sociology*, 110(2):349–399, 2004.

[Bur09] Ronald S Burt. *Structural holes: The social structure of competition*. Harvard university press, 2009.

[CDM+08] Peter Craig, Paul Dieppe, Sally Macintyre, Susan Michie, Irwin Nazareth, and Mark Petticrew. Developing and evaluating complex interventions: the new medical research council guidance. *BMJ*, 337, 2008.

[CDSLS93] Nancy L Collins, Christine Dunkel-Schetter, Marci Lobel, and Susan C Scrimshaw. Social support in pregnancy: Psychosocial correlates of birth outcomes and postpartum depression. *Journal of personality and social psychology*, 65(6):1243, 1993.

[CFSV06] Vittoria Colizza, Alessandro Flammini, M Angeles Serrano, and Alessandro Vespignani. Detecting rich-club ordering in complex networks. *Nature physics*, 2(2):110–115, 2006.

[CG88] Nikolas Coupland and Howard Giles. Introduction: The communicative contexts of accommodation. *Language & Communication*, 1988.

[CGHH12] Erik Cambria, Marco Grassi, Amir Hussain, and Catherine Havasi. Sentic computing for social media marketing. *Multimedia tools and applications*, 59(2):557–577, 2012.

[CK17] Munmun De Choudhury and Emre Kiciman. The language of social support in social media and its effect on suicidal ideation risk. In *Proceedings of the International Conference onÂ Web and Social Media (ICWSM)*. AAAI, May 2017.

[CLPS12] L Cecchi, G Liccardi, F Pellegrino, and M Sofia. 2 social networks: A new source of psychological stress or a way to enhance self-esteem? negative and positive implications in bronchial asthma. *Journal of Investigational Allergology and Clinical Immunology*, 22(6):402, 2012.

[Col88] James S Coleman. Social capital in the creation of human capital. *American journal of sociology*, 94:S95–S120, 1988.

[COR+16] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016.

[Cou05] Neil S Coulson. Receiving social support online: an analysis of a computer-mediated support group for individuals living with irritable bowel syndrome. *CyberPsychology & Behavior*, 8(6):580–584, 2005.

[Cra15] Matthew Crawford. *The World Beyond Your Head: How to Flourish in an Age of Distraction*. penguin publishing, 2015.

[CS92] Carolyn E Cutrona and Julie A Suhr. Controllability of stressful events and satisfaction with spouse support behaviors. *Communication Research*, 19(2):154–174, 1992.

[CSH+17] Daniel TC Cox, Danielle F Shanahan, Hannah L Hudson, Kate E Plummer, Gavin M Siriwardena, Richard A Fuller, Karen Anderson, Steven Hancock, and Kevin J Gaston. Doses of neighborhood nature: the benefits for mental health of living with nature. *BioScience*, 67(2):147–155, 2017.

[D+08] Ritendra Datta et al. Algorithmic inferencing of aesthetics and emotion in natural images: An exposition. In *2008 15th IEEE ICIP*, pages 105–108. IEEE, 2008.

[Dav67] James A Davis. Clustering and structural balance in graphs. *Human relations*, 20(2):181–187, 1967.

[DB01] Thomas H Davenport and John C Beck. *The attention economy: Understanding the new currency of business*. Harvard Business Press, 2001.

[DB08] A. De Botton. *The Architecture of Happiness*. Vintage Series. Knopf Doubleday Publishing Group, 2008.

[DB16] Alexey Dosovitskiy and Thomas Brox. Generating images with perceptual similarity metrics based on deep networks. In *Advances in Neural Information Processing Systems*, pages 658–666, 2016.

[DCKD+16] Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2098–2110, New York, NY, USA, 2016. ACM.

[DD14] Munmun De Choudhury and Sushovan De. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. *Proceedings of the Eight International AAAI Conference on Weblogs and Social Media*, pages 71–80, 2014.

[DiM04] M Robin DiMatteo. Social support and patient adherence to medical treatment: a meta-analysis. *Health psychology*, 23(2):207, 2004.

[DL67] James A Davis and Samuel Leinhardt. The structure of positive interpersonal relations in small groups. *Human relations*, 1967.

[DNP+16] Abhimanyu Dubey, Nikhil Naik, Devi Parikh, Ramesh Raskar, and César A Hidalgo. Deep learning the city: Quantifying urban perception at a global scale. *arXiv preprint arXiv:1608.01769*, 2016.

[DNVZ+16a] Marco De Nadai, Radu Laurentiu Vieriu, Gloria Zen, Stefan Dragicevic, Nikhil Naik, Michele Caraviello, Cesar Augusto Hidalgo, Nicu Sebe, and Bruno Lepri. Are Safer Looking Neighborhoods More Lively?: A Multimodal Investigation into Urban Life. In *Proceedings of the ACM on Multimedia Conference (MM)*, 2016.

[DNVZ+16b] Marco De Nadai, Radu Laurentiu Vieriu, Gloria Zen, Stefan Dragicevic, Nikhil Naik, Michele Caraviello, Cesar Augusto Hidalgo, Nicu Sebe, and Bruno Lepri. Are safer looking neighborhoods more lively?: A multimodal investigation into urban life. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 1127–1135. ACM, 2016.

[DS84] Christine Dunkel-Schetter. Social support and cancer: Findings based on patient interviews and their implications. *Journal of Social issues*, 40(4):77–98, 1984.

[DSHF+17] Anna De Simoni, Robert Horne, Louise Fleming, Andrew Bush, and Chris Griffiths. What do adolescents with asthma really think about adherence to inhalers? insights from a qualitative analysis of a uk online forum. *BMJ open*, 7(6):e015245, 2017.

[DSSBSM16] Anna De Simoni, Andrew Shanks, Chantal Balasooriya-Smeekens, and Jonathan Mant. Stroke survivors and their families receive information and support on an individual basis from an online forum: descriptive analysis of a population of 2348 patients and qualitative study of a sample of participants. *BMJ open*, 6(4):e010501, 2016.

[DVZK+14] Lance De Vine, Guido Zuccon, Bevan Koopman, Laurianne Sitbon, and Peter Bruza. Medical semantic similarity with a neural language model. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management*, pages 1819–1822. ACM, 2014.

[EC13] Reid Ewing and Otto Clemente. *Measuring Urban Design: Metrics for livable places*. Island Press, 2013.

[Eik17] Roelof Eikelboom. Irresistible: The rise of addictive technology and the business of keeping us hooked. *Perspectives on Science and Christian Faith*, 69(4):253–256, 2017.

[Fau07] Katherine Faust. 7. very local structure in social networks. *Sociological Methodology*, 37(1):209–256, 2007.

[Fau08] Katherine Faust. Triadic configurations in limited choice sociometric networks: Empirical and theoretical results. *Social Networks*, 30(4):273–282, 2008.

[FM08] Jeana Frost and Michael Massagli. Social uses of personal health information within patientslikeme, an online patient community: what can happen when patients have access to one another's data. *Journal of medical Internet research*, 10(3):e15, 2008.

[FMK+10] Tim Finin, Will Murnane, Anand Karandikar, Nicholas Keller, Justin Martineau, and Mark Dredze. Annotating named entities in twitter data with crowdsourcing. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 80–88. Association for Computational Linguistics, 2010.

[Fox11] Susannah Fox. *The social life of health information 2011*. 2011.

[GCBK+05] Billie Giles-Corti, Melissa H Broomhall, Matthew Knuiman, Catherine Collins, Kate Douglas, Kevin Ng, Andrea Lange, and Robert J Donovan. Increasing walking: how important is distance to, attractiveness, and size of public open space? *American Journal of Preventive Medicine*, 28(2):169–176, 2005.

[GCL09] Jeannette Golden, Ronán M Conroy, and Brian A Lawlor. Social support network structure in older people: underlying dimensions and association with psychological and physical health. *Psychology, Health & Medicine*, 14(3):280–290, 2009.

[GHS16] Erik Gawehn, Jan A Hiss, and Gisbert Schneider. Deep learning in drug discovery. *Molecular informatics*, 35(1):3–14, 2016.

[Gko17] George Gkotsis. Characterisation of mental health conditions in social media using informed deep learning. *Scientific Reports*, 7:45141, 2017.

[GPAM+14] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2014.

[Gra77] Mark S Granovetter. The strength of weak ties. In *Social networks*, pages 347–367. Elsevier, 1977.

[GTMM+15] Mireia Gascon, Margarita Triguero-Mas, David Martínez, Payam Dadvand, Joan Forns, Antoni Plasència, and Mark Nieuwenhuijsen. Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *International journal of environmental research and public health*, 12(4):4354–4379, 2015.

[HBCF16] William R Hobbs, Moira Burke, Nicholas A Christakis, and James H Fowler. Online social integration is associated with reduced mortality risk. *Proceedings of the National Academy of Sciences*, 113(46):12980–12984, 2016.

[HKR+18] Marco Helbich, Nadja Klein, Hannah Roberts, Paulien Hagedoorn, and Peter P Groenewegen. More green space is related to less antidepressant prescription rates in the netherlands: A bayesian geoadditive quantile regression approach. *Environmental research*, 166:290–297, 2018.

[HL71] Paul W. Holland and Samuel Leinhardt. Transitivity in structural models of small groups. *Comparative Group Studies*, 2(2):107–124, 1971.

[HL76] Paul W Holland and Samuel Leinhardt. Local structure in social networks. *Sociological methodology*, 7:1–45, 1976.

[HL77] Paul W Holland and Samuel Leinhardt. A method for detecting structure in sociometric data. In *Social Networks*, pages 411–432. Elsevier, 1977.

[HMG07] Ralf Herbrich, Tom Minka, and Thore Graepel. Trueskill: a bayesian skill rating system. In *Advances in Neural Information Processing Systems*, pages 569–576, 2007.

[HOG+10] Kevin O Hwang, Allison J Ottenbacher, Angela P Green, M Roseann Cannon-Diehl, Oneka Richardson, Elmer V Bernstam, and Eric J Thomas. Social support in an internet weight loss community. *International journal of medical informatics*, 79(1):5–13, 2010.

[How06] Jeff Howe. The rise of crowdsourcing. *Wired magazine*, 14(6):1–4, 2006.

[IABS+17] Nkeonye J Izuka, Matthew AW Alexander, Chantal Balasooriya-Smeekens, Jonathan Mant, and Anna De Simoni. How do stroke survivors and their carers use practitioners' advice on secondary prevention medications? qualitative study of an online forum. *Family practice*, 34(5):612–620, 2017.

[IKW09] Shereene Z Idriss, Joseph C Kvedar, and Alice J Watson. The role of online support communities: benefits of expanded social networks to patients with psoriasis. *Archives of dermatology*, 145(1):46–51, 2009.

[Jac61] J. Jacobs. *The Death and Life of Great American Cities*. Random House, 1961.

[JBS+11] Justin Jagosh, Joseph Donald Boudreau, Yvonne Steinert, Mary Ellen MacDonald, and Lois Ingram. The importance of physician listening from the patients' perspective: Enhancing diagnosis, healing, and the doctor–patient relationship. *Patient education and counseling*, 85(3):369–374, 2011.

[JK19] Kee Moon Jang and Youngchul Kim. Crowd-sourced cognitive mapping: A new way of displaying people's cognitive perception of urban space. *PloS one*, 14(6):e0218590, 2019.

[JSC+18] Sagar Joglekar, Nishanth Sastry, Neil S Coulson, Stephanie JC Taylor, Anita Patel, Robbie Duschinsky, Amrutha Anand, Matt Jameson Evans, Chris J Griffiths, Aziz Sheikh, et al. How online communities of people with long-term conditions function and evolve: Network analysis of the structure and dynamics of the asthma uk and british lung foundation online communities. *Journal of medical Internet research*, 20(7), 2018.

[JSR17] Sagar Joglekar, Nishanth Sastry, and Miriam Redi. Like at first sight: Understanding user engagement with the world of microvideos. In *International Conference on Social Informatics*, pages 237–256. Springer, 2017.

[K+12] Alex Krizhevsky et al. Imagenet classification with deep convolutional neural networks. In *NIPS*, pages 1097–1105, 2012.

[KAH+16] Christina Katsimerou, Joris Albeda, Alina Huldtgren, Ingrid Heynderickx, and Judith A Redi. Crowdsourcing empathetic intelligence: the case of the annotation of emma database for emotion and mood recognition. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 7(4):1–27, 2016.

[Kan87] Immanuel Kant. *Critique of judgment*. Hackett Publishing, 1987.

[KGB+02] Per E Kummervold, Deede Gammon, Svein Bergvik, Jan-Are K Johnsen, Toralf Hasvold, and Jan H Rosenvinge. Social support in a wired world:

use of online mental health forums in norway. *Nordic journal of psychiatry*, 56(1):59–65, 2002.

[KK89] Rachel Kaplan and Stephen Kaplan. *The experience of nature: A psychological perspective*. CUP Archive, 1989.

[KKW72] Stephen Kaplan, Rachel Kaplan, and John S Wendt. Rated preference and complexity for natural and urban visual material. *Perception & Psychophysics*, 12(4):354–356, 1972.

[KLP13] Yelin Kim, Honglak Lee, and Emily Mower Provost. Deep learning for robust feature generation in audiovisual emotion recognition. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 3687–3691. IEEE, 2013.

[KM16] Susan Kirk and Linda Milnes. An exploration of how young people and parents use online support in the context of living with cystic fibrosis. *Health Expectations*, 19(2):309–321, 2016.

[KRMH+16] Ramakanth Kavuluru, María Ramos-Morales, Tara Holaday, Amanda G. Williams, Laura Haye, and Julie Cerel. Classification of helpful comments on online suicide watch forums. In *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, BCB '16, pages 32–40, New York, NY, USA, 2016. ACM.

[KV16] Edward Kim and Shruthika Vangala. Vinereactor: Crowdsourced spontaneous facial expression data. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval*, pages 221–224, 2016.

[KW06] Gueorgi Kossinets and Duncan J Watts. Empirical analysis of an evolving social network. *science*, 311(5757):88–90, 2006.

[KZG+17] Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A Shamma, et al. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International Journal of Computer Vision*, 123(1):32–73, 2017.

[LCB01] Nan Lin, Karen S Cook, and Ronald S Burt. *Social capital: Theory and research*. Transaction Publishers, 2001.

[LH12] Pall Jakob Lindal and Terry Hartig. Architectural variation, building height, and the restorative quality of urban residential streetscapes. *Journal of Environmental Psychology*, 2012.

[LLCG17] Victor OK Li, Jacqueline CK Lam, Yun Chen, and Jiatao Gu. Deep learning model to estimate air pollution using m-bp to fill in missing proxy urban data. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.

[LM11] Andrew CK Lee and Ravi Maheswaran. The health benefits of urban green spaces: a review of the evidence. *Journal of public health*, 33(2):212–222, 2011.

[LM14] Quoc Le and Tomas Mikolov. Distributed representations of sentences and documents. In *International conference on machine learning*, pages 1188–1196, 2014.

[LPR18] Stephen Law, Brooks Paige, and Chris Russell. Take a look around: using street view and satellite images to estimate house prices. *arXiv preprint arXiv:1807.07155*, 2018.

[LPR19] Stephen Law, Brooks Paige, and Chris Russell. Take a look around: using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(5):1–19, 2019.

[LS12] Yung-Ming Li and Ya-Lin Shiu. A diffusion mechanism for social advertising over microblogs. *Decision Support Systems*, 54(1):9–22, 2012.

[LSSGR18] Stephen Law, Chanuki Illushka Seresinhe, Yao Shen, and Mario Gutierrez-Roig. Street-Frontage-Net: urban image classification using deep convolutional neural networks. *International Journal of Geographical Information Science*, pages 1–27, 2018.

[LTH+13] Tiancheng Lou, Jie Tang, John Hopcroft, Zhanpeng Fang, and Xiaowen Ding. Learning to predict reciprocity and triadic closure in social networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 7(2):5, 2013.

[LTH+17] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4681–4690, 2017.

[LYHC15] Kevin Lin, Huei-Fang Yang, Jen-Hao Hsiao, and Chu-Song Chen. Deep learning of binary hash codes for fast image retrieval. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2015.

[Lyn60] Kevin Lynch. *The image of the city*. 1960.

[May18] Susan Mayor. Nhs long term plan must tackle current challenges to set realistic targets, says nhs providers, 2018.

[MC12] Phoenix KH Mo and Neil S Coulson. Developing a model for online support group use, empowering processes and psychosocial outcomes for individuals living with hiv/aids. *Psychology & Health*, 27(4):445–459, 2012.

[MD03] Gordon Mitchell and Danny Dorling. An environmental justice analysis of british air quality. *Environment and planning A*, 35(5):909–929, 2003.

[MdFCC07] Julian J McAuley, Luciano da Fontoura Costa, and Tibério S Caetano. Rich-club phenomenon across complex network hierarchies. *Applied Physics Letters*, 91(8):084103, 2007.

[MMA+13] Lea Ann Matura, Annette McDonough, Lianne M Aglietti, Jessica L Herzog, and Kelly A Gallant. A virtual community: concerns of patients with pulmonary hypertension. *Clinical nursing research*, 22(2):155–171, 2013.

[MMG+07] Alan Mislove, Massimiliano Marcon, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 29–42. ACM, 2007.

[MN16] Todor Mihaylov and Preslav Nakov. Semanticz at semeval-2016 task 3: Ranking relevant answers in community question answering using semantic similarity based on fine-tuned word embeddings. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 879–886, 2016.

[Moo08] Guy Moors. Exploring the effect of a middle response category on response style in attitude measurement. *Quality and Quantity*, 42(6):779–794, 2008.

[MPP+13] Lev Muchnik, Sen Pei, Lucas C Parra, Saulo DS Reis, José S Andrade Jr, Shlomo Havlin, and Hernán A Makse. Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Scientific reports*, 3:1783, 2013.

[MSC+13] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.

[MSG+18] Liqian Ma, Qianru Sun, Stamatios Georgoulis, Luc Van Gool, Bernt Schiele, and Mario Fritz. Disentangled person image generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 99–108, 2018.

[MSGL14] Seth A Myers, Aneesh Sharma, Pankaj Gupta, and Jimmy Lin. Information network or social network?: the structure of the twitter follow graph. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 493–498. ACM, 2014.

[MSOI+02] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[MVF11] Gerald Mollenhorst, Beate Völker, and Henk Flap. Shared contexts and triadic closure in core discussion networks. *Social Networks*, 33(4):292–302, 2011.

[MVG+06] Jolanda Maas, Robert A Verheij, Peter P Groenewegen, Sjerp De Vries, and Peter Spreeuwenberg. Green space, urbanity, and health: how strong is the relation? *Journal of Epidemiology & Community Health*, 60(7):587–592, 2006.

[Nam11] Priya Nambisan. Information seeking and social support in online health communities: impact on patients' perceived empathy. *Journal of the American Medical Informatics Association*, 18(3):298–304, 2011.

[NDY+16] Anh Nguyen, Alexey Dosovitskiy, Jason Yosinski, Thomas Brox, and Jeff Clune. Synthesizing the preferred inputs for neurons in neural networks via deep generator networks. In *Advances in Neural Information Processing Systems*, pages 3387–3395, 2016.

[NKR+17] Nikhil Naik, Scott Duke Kominers, Ramesh Raskar, Edward L Glaeser, and César A Hidalgo. Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences*, 114(29):7571–7576, 2017.

[NPK13] D Nutsford, AL Pearson, and S Kingham. An ecological study investigating the association between access to urban green space and mental health. *Public health*, 127(11):1005–1011, 2013.

[NPRH14] Nikhil Naik, Jade Philipoom, Ramesh Raskar, and César Hidalgo. Streetscore-predicting the perceived safety of one million streetscapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pages 779–785, 2014.

[NWS02] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.

[OAS10] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social networks*, 32(3):245–251, 2010.

[OCPR08] Tore Opsahl, Vittoria Colizza, Pietro Panzarasa, and Jose J Ramasco. Prominence and control: the weighted rich-club effect. *Physical review letters*, 101(16):168702, 2008.

[Ops13] Tore Opsahl. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks*, 35(2):159–167, 2013.

[PB15] Pietro Panzarasa and Moreno Bonaventura. Emergence of long-range correlations and bursty activity patterns in online communication. *Physical Review E*, 92(6):062821, 2015.

[PC18] Albert Park and Mike Conway. Harnessing reddit to understand the written-communication challenges experienced by individuals with mental health disorders: Analysis of texts from mental health communities. *J Med Internet Res*, 20(4):e121, Apr 2018.

[PCC18] Albert Park, Mike Conway, and Annie T Chen. Examining thematic similarity, difference, and membership in three online mental health communities from reddit: a text mining and visualization approach. *Computers in human behavior*, 78:98–112, 2018.

[PCI10] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision making*, 5(5):411–419, 2010.

[Per14] Leonid Perlovsky. Aesthetic emotions, what are their cognitive functions? *Frontiers in Psychology*, 5:98, 2014.

[PFR+00] Robert S Peirce, Michael R Frone, Marcia Russell, M Lynne Cooper, and Pamela Mudar. A longitudinal model of social contact, social support, depression, and alcohol use. *Health Psychology*, 19(1):28, 2000.

[Pic03] Rosalind W Picard. Affective computing: challenges. *International Journal of Human-Computer Studies*, 59(1-2):55–64, 2003.

[PKD+16] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A Efros. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2536–2544, 2016.

[PPNH07] Maja Pantic, Alex Pentland, Anton Nijholt, and Thomas S Huang. Human computing and machine understanding of human behavior: A survey. In *Artifical Intelligence for Human Computing*, pages 47–71. Springer, 2007.

[PRS+14] Maria Panagioti, Gerry Richardson, Nicola Small, Elizabeth Murray, Anne Rogers, Anne Kennedy, Stanton Newman, and Peter Bower. Self-management support interventions to reduce health care utilisation without compromising outcomes: a systematic review and meta-analysis. *BMC health services research*, 14(1):356, 2014.

[PRT+16] Nikolaos Pappas, Miriam Redi, Mercan Topkara, Brendan Jou, Hongyi Liu, Tao Chen, and Shih-Fu Chang. Multilingual visual sentiment concept matching. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval*, pages 151–158, 2016.

[PS15] Louise F Pendry and Jessica Salvatore. Individual and social benefits of online discussion forums. *Computers in Human Behavior*, 50:211–220, 2015.

[PZ17] Jun Pang and Yang Zhang. Deepcity: A feature learning framework for mining location check-ins. In *Eleventh International AAAI Conference on Web and Social Media*, 2017.

[QASD15] Daniele Quercia, Luca Maria Aiello, Rossano Schifanella, and Adam Davies. The Digital Life of Walkable Streets. In *Proceedings of the 24th ACM Conference on World Wide Web (WWW)*, pages 875–884, 2015.

[QOC14] Daniele Quercia, Neil Keith O'Hare, and Henriette Cramer. Aesthetic capital: what makes london look beautiful, quiet, and happy? In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, pages 945–955, 2014.

[QPAC13] Daniele Quercia, Joao Paulo Pesce, Virgilio Almeida, and Jon Crowcroft. Psychological Maps 2.0: A web gamification enterprise starting in London. In *Proceedings of ACM International Conference on World Wide Web (WWW)*, 2013.

[QSA14] Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello. The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, pages 116–125, 2014.

[QSAM15] Daniele Quercia, Rossano Schifanella, Luca Maria Aiello, and Kate McLean. Smelly maps: the digital life of urban smellscapes. *arXiv preprint arXiv:1505.06851*, 2015.

[Que15] Daniele Quercia. Chatty, happy, and smelly maps. In *Proceedings of the 24th International Conference on World Wide Web*, pages 741–741. ACM, 2015.

[R+15] Miriam Redi et al. Like partying? your face says it all. predicting the ambiance of places with profile pictures. *arXiv preprint arXiv:1505.07522*, 2015.

[RAS00] Eulogio Real, Constantino Arce, and José Manuel Sabucedo. Classification of landscapes using quantitative and categorical data, and prediction of their scenic beauty in north-western spain. *Journal of environmental psychology*, 20(4):355–373, 2000.

[Row07] Jennifer Rowley. The wisdom hierarchy: representations of the dikw hierarchy. *Journal of information science*, 33(2):163–180, 2007.

[RPC17] Rajeev Ranjan, Vishal M Patel, and Rama Chellappa. Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):121–135, 2017.

[RWD+16] Daniele Ravì, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, and Guang-Zhong Yang. Deep learning for health informatics. *IEEE journal of biomedical and health informatics*, 21(1):4–21, 2016.

[SB84] Sally A Shumaker and Arlene Brownell. Toward a theory of social support: Closing conceptual gaps. *Journal of social issues*, 40(4):11–36, 1984.

[SBL+12] Shaun Scholes, Madhavi Bajekal, Hande Love, Nathaniel Hawkins, Rosalind Raine, Martin O'Flaherty, and Simon Capewell. Persistent socioeconomic inequalities in cardiovascular risk factors in england over 1994-2008: a time-trend analysis of repeated cross-sectional data. *BMC public health*, 12(1):129, 2012.

[SC16] Amie Shoebotham and Neil S Coulson. Therapeutic affordances of online support group use in women with endometriosis. *Journal of medical Internet research*, 18(5):e109, 2016.

[Sco88] John Scott. Social network analysis. *Sociology*, 22(1):109–127, 1988.

[SDP12] Yading Song, Simon Dixon, and Marcus Pearce. A survey of music recommendation systems and future perspectives. In *9th International Symposium on Computer Music Modeling and Retrieval*, volume 4, 2012.

[SFSJ06] Harald T Schupp, Tobias Flaisch, Jessica Stockburger, and Markus Junghöfer. Emotion and attention: event-related brain potential studies. *Progress in brain research*, 156:31–51, 2006.

[SHK+14] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

[SLJ+15] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.

[SLM+11] Miriam Stewart, Nicole Letourneau, Jeffrey R Masuda, Sharon Anderson, and Shawna McGhan. Online solutions to support needs and preferences of parents of children with asthma and allergies. *Journal of Family Nursing*, 17(3):357–379, 2011.

[SM12] Daizaburo Shizuka and David B McDonald. A social network perspective on measurements of dominance hierarchies. *Animal Behaviour*, 83(4):925–934, 2012.

[SM15] Daizaburo Shizuka and David B McDonald. The network motif architecture of dominance hierarchies. *Journal of the Royal Society Interface*, 12(105):20150080, 2015.

[Soa15] Molly Soat. Social media triggers a dopamine high. *Marketing News*, 49(11):20–21, 2015.

[Spe12] J. Speck. Walkable City: How Downtown Can Save America, One Step at a Time. In *Farrar, Straus and Giroux*, 2012.

[SPM15] Chanuki Illushka Seresinhe, Tobias Preis, and Helen Susannah Moat. Quantifying the impact of scenic environments on health. *Scientific Reports*, 5:16899, 2015.

[SPM17] Chanuki Illushka Seresinhe, Tobias Preis, and Helen Susannah Moat. Using deep learning to quantify the beauty of outdoor places. *Royal Society Open Science*, 4(7):170170, 2017.

[Squ15] Megan Squire. " should we move to stack overflow?" measuring the utility of social media for developer support. In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, volume 2, pages 219–228. IEEE, 2015.

[SR04] Robert J. Sampson and Stephen W. Raudenbush. Seeing Disorder: Neighborhood Stigma and the Social Construction of Broken Windows. *Social Psychology Quarterly*, 67(4), 2004.

[SSH13] Philip Salesses, Katja Schechtner, and César A Hidalgo. The collaborative image of the city: mapping the inequality of urban perception. *PloS one*, 8(7):e68400, 2013.

[ST14] Stavros Sintos and Panayiotis Tsaparas. Using strong triadic closure to characterize ties in social networks. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1466–1475. ACM, 2014.

[SY06] William P Sacco and Tovah Yanover. Diabetes and depression: the role of social support and medical symptoms. *Journal of behavioral medicine*, 29(6):523–531, 2006.

[TGL+14] Abdonas Tamosiunas, Regina Grazuleviciene, Dalia Luksiene, Audrius Dedele, Regina Reklaitiene, Migle Baceviciene, Jone Vencloviene, Gailute Bernotiene, Ricardas Radisauskas, Vilija Malinauskiene, et al. Accessibility and use of urban green spaces, and cardiovascular health: findings from a kaunas cohort study. *Environmental Health*, 13(1):20, 2014.

[TMM16] Gonçalo Tavares, André Mourão, and João Magalhães. Crowdsourcing facial expressions for affective-interaction. *Computer Vision and Image Understanding*, 147:102–113, 2016.

[TMP12] Amanda L Traud, Peter J Mucha, and Mason A Porter. Social structure of facebook networks. *Physica A: Statistical Mechanics and its Applications*, 391(16):4165–4180, 2012.

[TNW02] Takehito Takano, Keiko Nakamura, and Masafumi Watanabe. Urban residential environments and senior citizens' longevity in megacity areas: the importance of walkable green spaces. *Journal of Epidemiology & Community Health*, 56(12):913–918, 2002.

[TOS+06] Riitta Toivonen, Jukka-Pekka Onnela, Jari Saramäki, Jörkki Hyvönen, and Kimmo Kaski. A model for social networks. *Physica A: Statistical Mechanics and its Applications*, 371(2):851–860, 2006.

[TPE+14] Stephanie JC Taylor, Hilary Pinnock, Eleni Epiphaniou, Gemma Pearce, Hannah L Parke, Anna Schwappach, Neetha Purushotham, Sadhana Jacob, Chris J Griffiths, Trisha Greenhalgh, et al. A rapid synthesis of the evidence on interventions supporting self-management for people with long-term conditions: Prisms–practical systematic review of self-management support for long-term conditions. 2014.

[TPW16] Yaniv Taigman, Adam Polyak, and Lior Wolf. Unsupervised cross-domain image generation. *arXiv preprint arXiv:1611.02200*, 2016.

[Ulr83] Roger S Ulrich. Aesthetic and affective response to natural environment. In *Behavior and the Natural Environment*, pages 85–125. Springer, 1983.

[Vau08] Brandon K Vaughn. Data analysis using regression and multilevel/hierarchical models. *Journal of Educational Measurement*, 45(1):94–97, 2008.

[VPR13] Carl Vondrick, Donald Patterson, and Deva Ramanan. Efficiently scaling up crowdsourced video annotation. *International journal of computer vision*, 101(1):184–204, 2013.

[VQC+15] Alessandro Venerandi, Giovanni Quattrone, Licia Capra, Daniele Quercia, and Diego Saez-Trumper. Measuring urban deprivation from user generated content. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 254–264. ACM, 2015.

[VS16] Domonkos Varga and Tamás Szirányi. Fast content-based image retrieval using convolutional neural network and hash function. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 002636–002640. IEEE, 2016.

[WB13] Helle S Wentzer and Ann Bygholm. Narratives of empowerment and compliance: studies of communication in online patient support groups. *International journal of medical informatics*, 82(12):e386–e394, 2013.

[WLH14] Cheng Wang, Omar Lizardo, and David S Hachen. Triadic evolution in a large-scale mobile phone network. *Journal of Complex Networks*, 3(2):264–290, 2014.

[WRP+94] Donald E Woodhouse, Richard B Rothenberg, John J Potterat, William W Darrow, Stephen Q Muth, Alden S Klovdahl, Helen P Zimmerman, Helen L Rogers, Tammy S Maldonado, John B Muth, et al. Mapping a social network of heterosexuals at high risk for hiv infection. *Aids*, 8(9):1331–1336, 1994.

[WW09] Richard TA Wood and Sabrina A Wood. An evaluation of two united kingdom online support forums designed to help people with gambling issues. *Journal of Gambling Issues*, (23):5–30, 2009.

[Yal95] Irvin D. Yalom. *The theory and practice of group psychotherapy, 4th ed.* The theory and practice of group psychotherapy, 4th ed. Basic Books, New York, NY, US, 1995.

[YLSK+04] Esti Yeger-Lotem, Shmuel Sattath, Nadav Kashtan, Shalev Itzkovitz, Ron Milo, Ron Y Pinter, Uri Alon, and Hanah Margalit. Network motifs in integrated cellular networks of transcription–regulation and protein–protein interaction. *Proceedings of the National Academy of Sciences*, 101(16):5934–5939, 2004.

[YYSL15] Xinchen Yan, Jimei Yang, Kihyuk Sohn, and Honglak Lee. Attribute2image: Conditional image generation from visual attributes. corr abs/1512.00570, 2015.

[ZAA07] Jun Zhang, Mark S Ackerman, and Lada Adamic. Expertise networks in online communities: structure and algorithms. In *Proceedings of the 16th international conference on World Wide Web*, pages 221–230. ACM, 2007.

[ZC11] Jonathan R Zadra and Gerald L Clore. Emotion and perception: The role of affective information. *Wiley interdisciplinary reviews: cognitive science*, 2(6):676–685, 2011.

[ZLX+14] Bolei Zhou, Agata Lapedriza, Jianxiong Xiao, Antonio Torralba, and Aude Oliva. Learning deep features for scene recognition using places database. In *Advances in Neural Information Processing Systems*, pages 487–495, 2014.

[ZM04] Shi Zhou and Raúl J Mondragón. The rich-club phenomenon in the internet topology. *IEEE Communications Letters*, 8(3):180–182, 2004.

[ZSSH14] Xin Zhang, Shuai Shao, H Eugene Stanley, and Shlomo Havlin. Dynamic motifs in socio-economic networks. *EPL (Europhysics Letters)*, 108(5):58001, 2014.

[ZYW17] Yongjun Zhu, Erjia Yan, and Fei Wang. Semantic relatedness and similarity of biomedical terms: examining the effects of recency, size, and section of biomedical publications on the performance of word2vec. *BMC medical informatics and decision making*, 17(1):95, 2017.

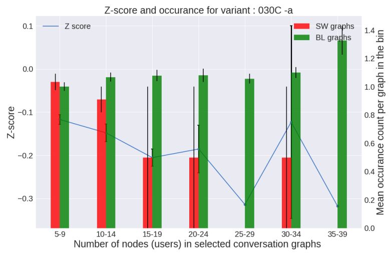
APPENDIX A

APPENDIX 1: RARE ANCHORED MOTIFS IN SOCIAL SUPPORT

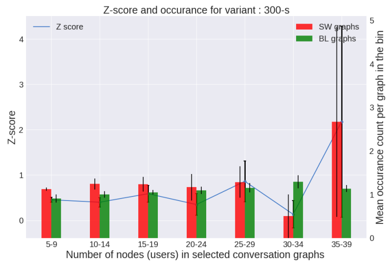
I evaluate the prevalence of all the Anchored Triadic motifs as defined in Chapter 4. As explored by previous studies [SM15, SM12, HL76, HL71], the social-tie structures and certain triadic motifs are predictors of social hierarchy in communities. Albeit these studies looked at community structures of Apes and tribal populations, they act as a very peculiar proxies for understanding how communities in the wild evolve. The important point to note here is that graphs that emerge from conversations, may not be synonymous to the graphs that form due to actual social ties. But in the context of a conversation, the interactions between peers can be assumed to be a purely transactional tie.

The internet has opened up new mechanisms of formation of these ties. And at that, the mechanisms seem to generate new patterns of conversations, based on the kind of interaction the participants of a community perform. To that end both the supportive and baseline communities exhibits a dearth of transitive triads like *030T* and *210C*, which are more prevalent in real life communities with hierarchies

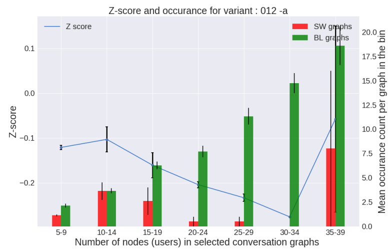
One of the most interesting outcome of this exercises, which I failed to pursue further is the complete lack of triadic closures. Triadic closure has been a very important aspect of studies around online social capital. The key premise here is that a weak tie can act as a bridge between two disjoint nodes in a community. However, in the context of conversation structures, this phenomenon seems to be completely absent from the picture.



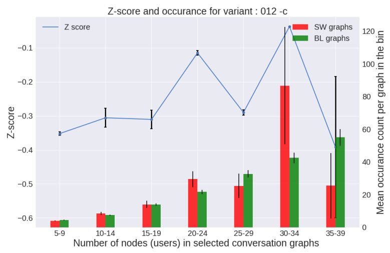
(1.a)



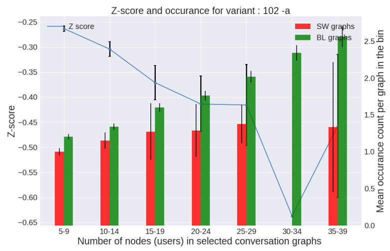
(1.b)



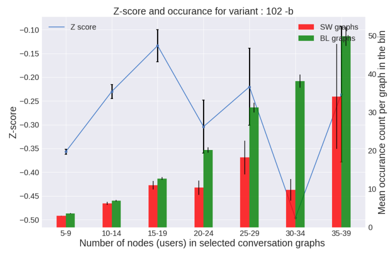
(1.c)



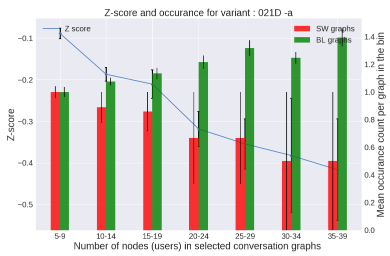
(2.a)



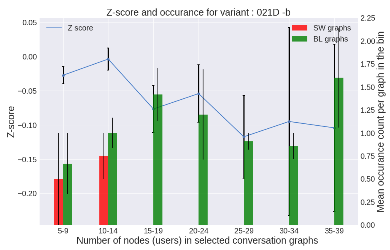
(2.b)



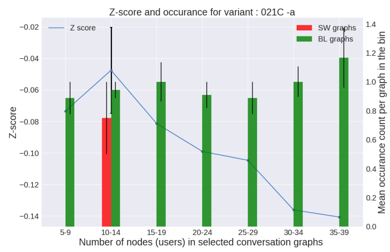
(2.c)



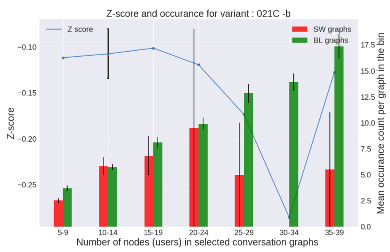
(3.a)



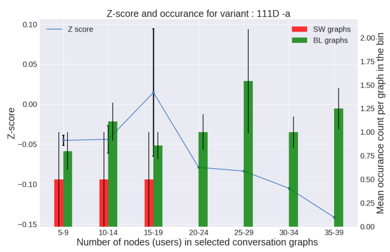
(3.b)



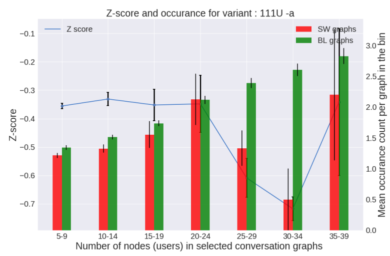
(3.c)



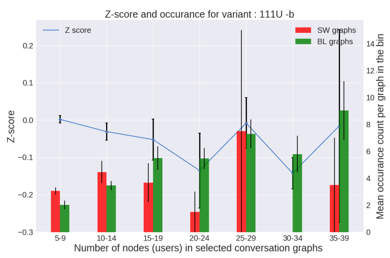
(4.a)



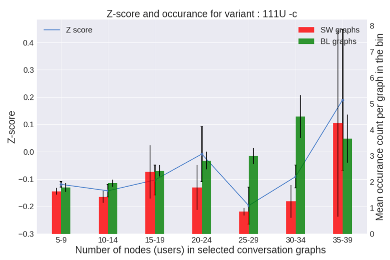
(4.b)



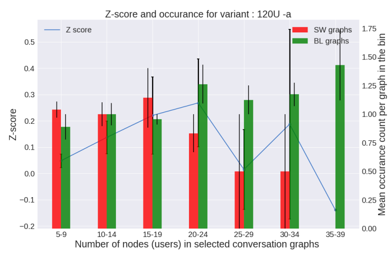
(4.c)



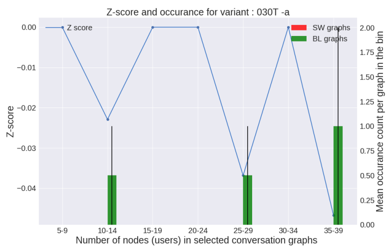
(5.a)



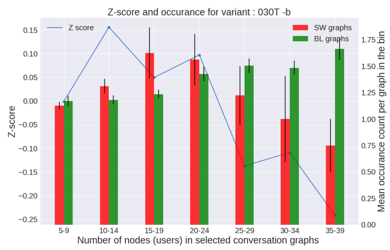
(5.b)



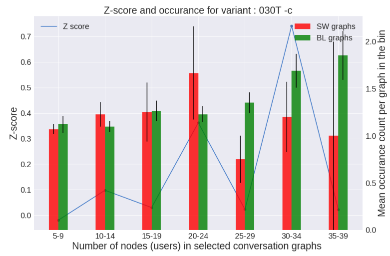
(5.c)



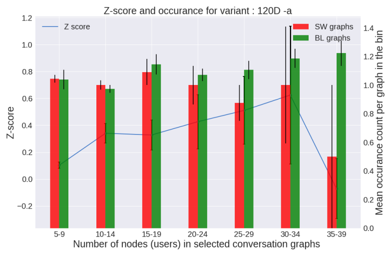
(6.d)



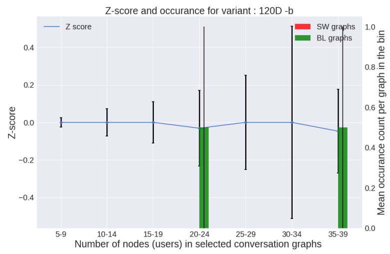
(6.e)



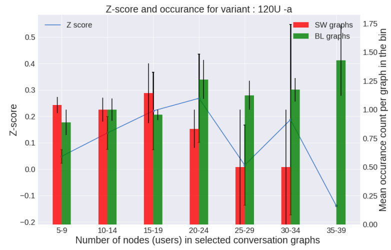
(6.f)



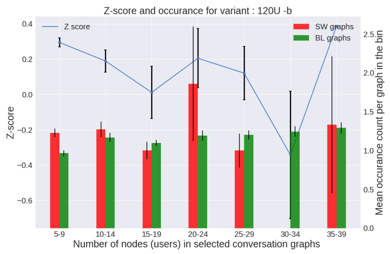
(7.a)



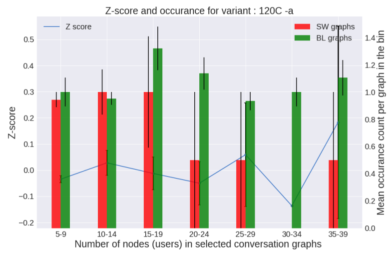
(7.b)



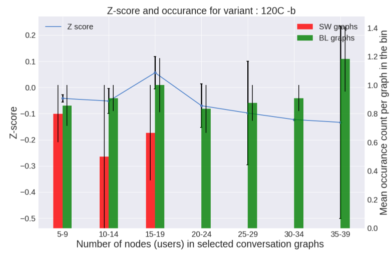
(7.c)



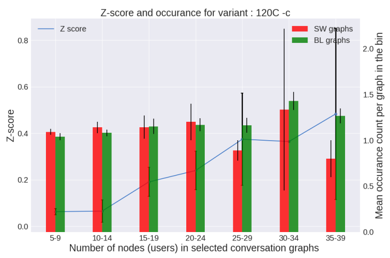
(8.a)



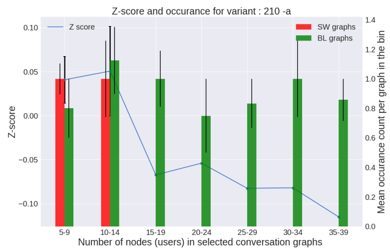
(8.b)



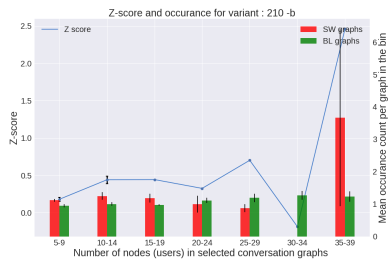
(8.c)



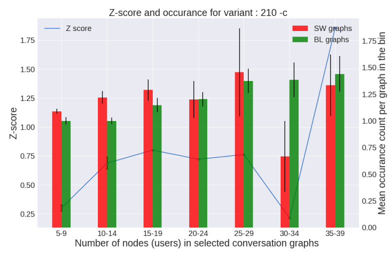
(9.a)



(9.b)



(9.c)



(10.a)

Fig. A.0 This figure lists out all the insignificant Anchored motifs, either by the virtue of rare occurance (<5 mean motifs per bin) or by account of low Z-score.

Appendix B

Appendix 2: Categorization of PlacesNet labels

As discussed in Chapter 6 I classified the labels from Placesnet, a deep learning framework to recognize the types of outdoor places [ZLX+14], into 4 distinct categories namely Architectural, Natural, Landmarks and Walkable. These categories were picked up to understand the effect of 4 broad classes of built and natural entities on urban beauty. The definitions for the category descriptions were defined based on the guidelines from [EC13].

Definition 8 ***Walkable*** *A scene from PlacesNet was defined to be walkable, if the scene is perceived to facilitate people to walk.*

Definition 9 ***Landmarks*** *A scene from PlacesNet was defined to be a Landmark, if the scene can be used to articulate, remember or communicate location relative to an object.*

Definition 10 ***Architectural*** *A scene from PlacesNet was defined to be Architectural, if the scene can be described to be a part of the city's built environment.*

Definition 11 ***Natural*** *A scene from PlacesNet was defined to be Natural, if the scene evokes the feeling about being in proximity with nature.*

Once the labels were classified, I developed metrics around the built environment of a place, based on the frequencies of labels belonging to the four categories. A detailed categorization of the scene labels into the four categories can be found in Table B.1.

Architectural	Walkable	Landmark	Natural
Apartment building	Abbey	Airport	Badlands
Building Facade	Alley	Amphitheatre	Bamboo Forest
Construction Site	Boardwalk	Amusement Park	Canyon
Courthouse	Botanical Garden	Arch	Coast
Drive way	Corridor	Amphitheatre	Corn field
Door way	Cottage garden	Baseball Field	Creek
Forest road	Courtyard	Basilica	Desert (Sand)
Garbage dump	Crosswalk	Bridge	Field (cultivated)
Golf course	Fairway	Castle	Field (wild)
Highway	Food court	Cemetery	Mountain
Hotel	Forest path	Cathedral	Snowy Mountain
Inn	Formal Garden	Church	Ocean
Ice skating rink	Herb Garden	Dam	Orchard
Motel	Outdoor Market	Dock	Pond
Office building	Nursery	Cemetery	Rainforest
Parking Lot	Patio	Fire station	Rice paddy
Railroad track	Pavilion	Fountain	River
Residential neighbourhood	Picnic area	Gas Station	Rock arch
Restaurant	Playground	Harbour	Sand bar
Runway	Plaza	Hospital	Sea Cliff
School House	Patio	Lighthouse	Ski slope
Skyscraper	Shopfront	Mansion	Sky
Slum	Topiary garden	Mausoleum	Snow field
Supermarket	Tree farm	Pagoda	Swamp
Outdoor swimming pool	Veranda	Palace	Valley
Tower	Vegetable garden	Racecourse	Wheat field
Water tower	Yard	Ruin	Desert (vegetation)
Wind farm		Rope Bridge	
		Ski Resort	
		Baseball stadium	
		Football stadium	
		Subway Station	
		Train Station	
		Temple	
		Wind mill	

Table B.1 Classification of Places-net labels into the four categories.