ASSIGNMENT COVERSHEET

UTS: ENGINEERING & INFORMATION TECHNOLOGY

SUBJECT NUMBER & NAME	NAME OF STUDENT(s) (PRINT CLEARLY)	STUDENT ID(s)
32144 - Technology Research Preparation	*Jared Mayger* *SURNAME* *FIRST NAME*	13887967

STUDENT EMAIL	STUDENT CONTACT NUMBER
jared.s.mayger@student.uts.edu.au	

NAME OF TUTOR	TUTORIAL GROUP	DUE DATE
Thomas Dolmark	**06**	**28-10-2022**

ASSESSMENT ITEM NUMBER & TITLE
Assessment Task 3 - Project Proposal

I confirm that I have read, understood and followed the guidelines for assignment submission and presentation on page 2 of this cover sheet.
I confirm that I have read, understood and followed the advice in the Subject Outline about assessment requirements.
I understand that if this assignment is submitted after the due date it may incur a penalty for lateness unless I have previously had an extension of time approved and have attached the written confirmation of this extension.

Declaration of originality: The work contained in this assignment, other than that specifically attributed to another source, is that of the author(s) and has not been previously submitted for assessment. I understand that, should this declaration be found to be false, disciplinary action could be taken and penalties imposed in accordance with University policy and rules. In the statement below, I have indicated the extent to which I have collaborated with others, whom I have named.

Statement of collaboration:

Signature of student(s) _____JM_____ Date _____27/10/2022_____

ASSIGNMENT RECEIPT

To be completed by the student if a receipt is required

SUBJECT NUMBER & NAME	NAME OF TUTOR

SIGNATURE OF TUTOR	RECEIVED DATE

Table of Contents

Executive Summary

The main purpose of this report is to outline how the maximisation in efficiency of data collection methods like text summarisation, image processing and topic modelling can be applied to e-commerce in social networks by the creation and mapping of customer personas, allowing the prediction of shopping interests and fashion tendencies. Significant research into the history and the development of existing data collection models has been undertaken for the fields of text summarisation, image processing, topic modelling and topic fusion. Motivation for this study stems from a business standpoint and the ability to market this technology to brands attempting to increase the conversion rate between social media targeted advertisements and product/service purchase. The main challenge in the increase ofc this conversion rate is the accuracy and efficency of data collection models, in coupled with the lack of coherence and redundancies found in topic mining from multimodal source as a result of the complicated nature of image topics. Throughout the study, data will be collected from social media sites like Intagram or Twitter, which will then be analysed using machine learning and artificial intelligence to produce appropriate topics pertaining to an individual customer or customers from a specific geographic or demographic. This evaluated information will be used to determine customer personas, successfully predict potential shopping interests of the individual and hopefully analyse fashion trends.

Indroduction

The lack of efficiency and accuracy in data collection and multimodal topic modelling in the context of e-commerce for social networks, leaves much to be desired when creating customer personas and targeting advertisements. Data collection methods such as text summarisation and adaptive nonparametric image parsing are combined and adjusted for multimodal sources such as social media sites like Instagram and Twitter. This paper aims to study the efficiency of existing models and where applicable, suggesting improvements that increase the effectiveness of the end goal. Most methods are already at an acceptable level of efficiency individually, however the inaccuracies come when correlating them and using them to work towards a common goal and hence similar topics. In the cases of small social media documents text is almost always directly related to the connected image, which in turn can be relevant to multiple topics due to the rich information expressed in the images, leading to low-quality topics (Zhang et al, 2022). Increasing the accuracy of these topic modelling algorithms will in turn allow customer personas, created by advertisers to more accurately represent the consumers wants and needs, proving extremely valuable to marketers and big corporations. An argument can also be made for government usefulness in terms of safely campaigns for example. This report will first detail the multiple aims of the research and the concurrent objectives, before providing a detailed background of the subject matter and the significance and field innovation of the solutions. Will finish up describing and justifying the methodologies that are most appropriate to throw light onto the topic of social network based multimodal topic modelling for e-commerce.

Research Aims and Objectives

Overall, the aim of this research is to study efficient data collection models, text and image based multimodal topic modelling, and its relation to e-commerce in social networks, specifically Twitter and Instagram. Objectively the main groups of research can be broken down into the following; efficient data collection, efficienct multimodal topic modelling and the accurate creation of customer personas. Existing text summarisation methods and adaptive nonparametric image processing techniques can be evaluated and hopefully improved for efficient data collection. Furthermore, combining the previous data collection methods with multimodal topic modelling methods for small area documents, like social media posts, we can hopefully see an improvement in the accuracy of the fused topics gathered by the algorithms. Using a all of the above the final objective is to collate speific topics regarding an individual and their online habits to more accurately predict their wants and needs and hence show them more personalized advertisements. This includes creating customer personas that can eventually be autonomous, and using them to persuade shopping tendencies with more success. It is expected that the fused topics will comprehensively reflect the customers' potential shopping interests, which are supported by the created customer persona and influences product recommendations. In addition, the research could be applied to fashion trend analysis to perhaps predict future trends, or even adapted to make other areas of consumer-advertisment interaction more effective and efficient.

Research Background

Text summarisation, specifically in its automatic form has been around since the 1950's (Luhn, 1958), however human summarisaiton undoubtedly dates back further as its usefulness in teaching and retaining information is paramount. An aspect of automatic summarisation is abstractive summarisation which plans to generate short summaries of essential information from long documents. Current models suffer from crucial problems like hallucinations (Maynez et al., 2020), low quality, factuality degradation and decreased efficiency, where summaries contain facts or entities not present in the original document. (Bansal et al, 2022). The FactPEGASUS model proposes a trident of components; a corrector to remove hallucinations, a contrastor to better differentiate between fact and non-factual summaries and a connector to bridge the gap between finetuning and pre-training (Bansal et al, 2022). The method was tested on three abstractive summarisation datasets and was able to succesfully improve factuality and conclude that the increase in extractivness is not solely responsible for overall factuality improvement.

In the case of customer reviews and other shorter, less connected writings like social media posts, most existing pre-trained models are not properly calibrated, leaving summaries full of semantic mistakes (Dreyer et al, 2022). As an alternative to fine-tuning the entirety of a model, adapters can be pre-trained task specifically and added onto a large dataset of unanotated reviews. Using human-annotations as a fine-tuning mechanism for adapters, becomes time consuming but is necessary for for the improvement of summary quality and can be utilized on smaller datasets. Pre-training adaters in a query based manner on customer reviews also results in better-organised summary content, improved coherence and fewer redundancies (Dreyer et al, 2022).

Significant exploration is needed into the collaboration between humans and artificial intelligence in order to determine whether post-editing summaries offers advantages in efficiency and the reduction of human workload. Conducted experiments with a fairly small group of participants found there were both benefits and drawbacks as participants needed assistance which became time consuming and almost all of them employed differing editing strategies (Jaimes et al, 2022). While, AI's attention for detail is lack luster, its efficiency and the standardised nature in which it summarises a text is unprecedented. In longer resources, where computational potential is decreased the quadratic memory complexity prevents long document summarisation. Additionally, models apply input truncation, resulting in a performance drop as potential summary-relevant content is left out, causing destruction for semantic text analysis (Ragazzi et al, 2022).'

An unsupervised extractive summarisation approach was able to successfully reduce topic bais by combining Latent Dirichlet Allocaiton for topic modelling and K-Medoids clustering (Kumar et al, 2022). One particular experiment made use of the Recall Oriented-Understudy for Gisting Evaluation (ROUGE) metrics for comparative analysis against recently reported techniques, specifically ROUGE-1 (R-1), ROUGE-2 (R-2), and ROUGE-L (R-L). The suggested framework offered scores of 34.80%, 9.13%, and 32.30% on the Wikihow Dataset, 43.90%, 19.01%, and 41.50% on the CNN/DailyMail Dataset, and 49.35%, 31.53%, and 41.72% on the DUC2002 Corpus (R-1, R-2, R-L respectively) (Kumar et al, 2022). Additionally other experiments suggested that topic modelling algorithms such as NonNegative Matrix Factorization and LSA could undoubtedly be beneficial in extractive text summarisation. (Salama et al, 2021).

Image parsing holds implicit challenges as it combines tasks such as object detection, segmentation, and multilabel recognition to which algorithmic solutions can be divided into two categories, paramentric and nonparametric. Many approaches to this problem have been proposed recently, includingpixel by pixel estimation(He et al. 2004; Ladicky et al. 2010; Shotton et al. 2006, 2008), ones that aggregate features over segmentation regions (Galleguillos et al. 2010; Gould et al. 2009; Hoiem et al. 2007; Malisiewicz and Efros 2008; Rabinovich et al. 2007; Socher et al. 2011), and others that successfully predict object bounding boxes (Divvala et al. 2009; Felzenszwalb et al. 2008; Heitz and Koller 2008; Russell et al. 2007).

Superpasing combines global scene-level matching against a training set followed by superpixel-level matching and efficient Markov random field (MRF) optimization for incorporating neighborhood context, is able to compute the labelling of image regions into geometric and semantic classes. The proposed method outperformed a state-of-the-art nonparametric method based on SIFT Flow on a dataset of 2,688 images and 33 labels (Lazebnik et al, 2012). In order to boost the superpixel matching process, locality-sensitive hashing (LSH) can be embedded to encode the features representative in few bits (instead of bytes) for large-scale matching (Yan et al, 2015). The prevailing consensus in the community is that image parsing requires context (Divvala et al. 2009; Galleguillos and Belongie 2010; Heitz and Koller 2008; Hoiem et al. 2007; Rabinovich et al. 2007), and its usefulness can ultimately applied to the research conducted in topic modeling and e-commerce in social media.

Regularized topic models are proved to be able to increase the coherence of learned topics when compared to the baseline LDA method, as assessed by human workers in Amazon Mechanical Turk (Wray et al, 2011). Social media posts are more multifaceted than simply text summarisation or image processing and it can be deduced that the text of a post is usually linked to one topic whilst the image is mapped to the same topic among others. As a result of the rich informaiton expressed in images (Zhang et al, 2022) a small number of cases see the topics of text and image to be different, hence past topic models fail to model these characteristics amd produce low-quality topics. An experiment conducted by (Zhang et al, 2022) propose an unsupervised multimodal topic model SMMTM to model the social media posts. In the SMMTM model, only one topic is sampled for the the text while an image can belong to different topics, as because of of the lack of co-ocurrence patterns, the assumption of multiple topics for short texts leads to data sparsity and an incoherance between topics. Other topic models like multimodal-LDA assume that there is only one topic for both the textual words and the visuals which are the same (Zhang et al, 2022), meaning that the new comprehensive SMMTM model was able to out-preform existing models.

Research Significance and Innovation

The particular significance of this study lies in its perceived ability to further improve data collection and topic modelling methods, and e-commerce approaches to social networks. We argue that any significant increase in the efficiency of data collection methods in the fields of text summarisation and image processing, which display direct correlation to the efficiency of topic modelling approaches will increase success in the creation of customer personas. Building on the improved factuality produced by the FactPEGASUS' text summarisation model on certain abstractive datasets (Bansal et al, 2022) in combination with further exploration into post-editing summarisation techniques (Jaimes et al, 2022), text summarisation as a whole should see efficiency increases, specifically for short text documents (ie; social media posts) (Dreyer et al, 2022). Furthermore by combining multiple non parapetric image processing algorithmic solutions to maximise object detection, segmentation, and multilabel recognition, hypothetically the ease in which topic keywords can be determined becomes less consuming. If it is determined that both text summarisation and image processing are efficient enough, more importance will be placed on multimodal topic modelling and converting learned topics into customer personas that can be used by advertisers to target more appropriate consumers in the hopes of increasing sales, goodwill, market share…etc. Existing multimodal topic models like the SMMTM model separate the topics collected from images and their corresponding texts, meaning accuracy is improved and the problems of data sparsity and inconsistency found when assuming multiple topics for short text documents is voided (Zhang et al, 2022).

Stakeholders of this research, both primary and secondary bridge the gap between numerous industries from tech and artificial intelligence to advertising and e-commerce and is mainly intriguing from a business point of view. In the technological sector the main stakeholders are analysts and developers who would be interested in different respective areas of the proposed research. Developers may be able to use that dat collection methods proposed to make further improvements in efficiency of data collection methods or even use the existing data collection for other projects. Similarly the data collection models could be adapted by data analysts for other research that require the autonomous collection of large quantities of data. Analysts would also be interested in the research as a result of the proposed increase in accuracy when portraying customer personas and predicting customer wants and needs. Further analysis could mean further improvements to the targeted advertisements that fuels the e-commerce economy. Inherently consumers themselves have some degree of vested interest in the technology and results of the research, if only on an awareness level, as they pertain directly to the personalised nature of the ads potential customers receive when browsing social networks. Finally the stakeholder who we believe will have the most interestin the the research and the means to pay for it are the brands themselves, smaller business that are able to guarantee a higher ads to purchase ratio will undoubtedly take that risk to grow and reach more customers with matching interests to their product.

Overall, most of this proposed project is an extension of already existing research and solutions, however will provide corporations with the novel opportunity to have more accurately targeted advertising. As a result of the more accurate creation of customer personas, detailed the potential wants of a client the marketing departments of companies

will be able to push ads that interest the consumer and increase sales. The research could be developed further to include a way to visualise these personas, and recommend which products to push to the consumer, however this will not be included in the scope of this research project.

Research Methods

In order to properly investigate the research area that is social network based multimodal topic modelling for e-commerce, a multitude of activities will be undertaken in a flexible research design format. This allows for more freedom throughout the data collection process and allows for the qualitative variables of interest, one of which is fashion trend analysis. The list of activities that we will aim to complete include but aren't limited to; efficient short document text data collection, adaptive nonparamteric image collection and processing, multimodal data collection using a web crawler, analysis of data using AI and machine learning, topic modelling and topic fusion. Topic modelling will allow us to identify keywords that will relate to a customer persona and reflect their potential shopping interests as well as introduce the concepts of a fashion trend and open the field up for analysis. We expect to collect and formulate a range of both quantitative and qualitative data pertaining to the efficiency of current and proposed data collection methods and the accuracy of fused topics in predictic a consumers shopping interests and future necessities. Qualitative observations will be made from existing research and presented in the form of a literature review, outlining the highlights and drawbacks of each method. Quantitatively, the efficiency of each text summarisation, image processing and topic modelling method can be compared against existing databases, human and artificial intelligence sources. The efficiency can be analyised using computer simulation and further computer analysis will be completed on the data collected from the resulting 'most efficient' methods. Since the majority of the data collected will be qualitative, let it be known that current sources and models need to be examined against recent datasets as the value of data decreases over time in a qualitative setting.

The overall apporach to this project is a combination of field study and action research. Comparative analysis between existing data collection models will be combined with the experimentation of siad models on new datasets and in multimodal areas. It is expected that these models and any proposed ones will be evaluated according to a specific criteria of accuracy adn efficiency where links can be made between the collected data and customer personas, aiding e-commerce in social networks. Measurements of performance will be recounted and picked apart to determine the efficiency of all data collection and topic modelling methods, providing a quantitative way to measure the study's progress. The accurate prediction of personal shopping interests and creation of customer personas, act as a benchmark for success within the project, with further acheivement coming from the increase in the efficiency of the process.

One particularly large limitation that we expect to encounter refers to both the sampling techniques and data collection strategy. As a result of the insanely large pool of social media posts daily, data collection for the purposes we intend, using a web crawler to scrape medias from a specific hashtag, would make the creation of individual customer personas near impossible. To address such problems, the crawled information could be altered to only

collect media from one particular user, however this would then highlight the time and resource limitations in conjunction with making the dataset extremely small. Post collection sorting of the data could be applied to possibly predict fashion trends in relation to a specific geographical location, meaning more generic customer personas would be created depending on geographic and demographic factors. In terms of accessibility, first-hand collection of media dat from well known social medai sites like Intagram, could pose security risks as they undoubtedly have anti-crawl programs in places. Existing datasets using the public Instagram API could be unitlised in the event of an IP block from the site or a lack of useable data collected. The main assumption of this study is the fact that without any way of proving so, the topics that are mined from multimodal data collected from social networks will explicitly link to the shopping interests of consumers and predict future fashion trends, allowing for more effective targeted advertisements. Undeniably, we cannot be sure of this connection, however due to the nature of present day social media we can deduce the prevalence of the connection.

Conclusions

In summation, this proposal aims to highlight the necessity of further research into social network based multimodal topic modelling in e-commerce and successfully address the problem being the lack of efficiency in data collection and topic modelling pertaining to social media and the creation of customer personas to expose potential shopping interests. We propose significant innovation in the maximization of text summarisation, image processing and topic mining models in conjunction with an apparent significance to stakeholders, like developers, analysts and brands that from a business standpoint could gain the ability to apply this research to their own initiatives. Existing models that are discussed in the research background, including the FactPEGASUS text summarisation model and more importantly the SMMTM methdo for topic modeling are a perfect starting point for the collection of multimodal topic data from short text social media documents. Furthermore, using a range of qualitative and quantitative research techniques (ie; most efficient data collection models) the successful fusion of collected topics and appropriate mapping of said topics into accurate customer personas becomes a more reliable system. With all this in mind, it is the ultimate goal of this study to successfuly predict potential customer shopping interests and analyise current fashion trends through the creation of accurate customer personas.

Reference List

1. Wan, & Bansal, M. (2022). FactPEGASUS: Factuality-Aware Pre-training and Fine-tuning for Abstractive Summarization. https://arxiv.org/pdf/2205.07830.pdf
2. Lai, Smith-Renner, A., Zhang, K., Cheng, R., Zhang, W., Tetreault, J., & Jaimes, A. (2022). An Exploration of Post-Editing Effectiveness in Text Summarization.https://arxiv.org/pdf/2206.06383.pdf
3. Bražinskas, Nallapati, R., Bansal, M., & Dreyer, M. (2022). Efficient Few-Shot Fine-Tuning for Opinion Summarization. https://arxiv.org/pdf/2205.02170.pdf
4. Xu, Zhang, X., Wu, Y., & Wei, F. (2021). Sequence Level Contrastive Learning for Text Summarization. https://www.aaai.org/AAAI22Papers/AAAI-7314.XuS.pdf
5. Srivastava, Singh, P., Rana, K. P. S., & Kumar, V. (2022). A topic modeled unsupervised approach to single document extractive text summarization. Knowledge-Based Systems, 246, 108636–. https://doi.org/10.1016/j.knosys.2022.108636
6. Liu, Shin, H., & Burns, A. C. (2021). Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing. Journal of Business Research, 125, 815–826. https://doi.org/10.1016/j.jbusres.2019.04.042
7. Chang, Trappey, A. J. C., Trappey, C. V., & Chien, J. Y. C. C. (2022). Intelligent RFQ Summarization Using Natural Language Processing, Text Mining, and Machine Learning Techniques. Journal of Global Information Management, 30(1), 1–26. https://doi.org/10.4018/JGIM.309082
8. Tighe, & Lazebnik, S. (2012). Superparsing: Scalable Nonparametric Image Parsing with Superpixels. International Journal of Computer Vision, 101(2), 329–349. https://doi.org/10.1007/s11263-012-0574-z
9. Nguyen, Lu, C., Sepulveda, J., & Yan, S. (2015). Adaptive Nonparametric Image Parsing. IEEE Transactions on Circuits and Systems for Video Technology, 25(10), 1565–1575. https://doi.org/10.1109/TCSVT.2014.2382982
10. Park, & Kim, Y.-G. (2021). Survey and challenges of story generation models - A multimodal perspective with five steps: Data embedding, topic modeling, storyline generation, draft story generation, and story evaluation. Information Fusion, 67, 41–63. https://doi.org/10.1016/j.inffus.2020.10.009
11. H. Zhang, C. Yi, B. Zhu, H. Ren and Q. Li, "Multimodal Topic Modeling by Exploring Characteristics of Short Text Social Media," in IEEE Transactions on Multimedia, doi: 10.1109/TMM.2022.3147064.
12. Newman, D. B., Edwin, Buntine, Wray. (2011). Improving Topic Coherence with Regularized Topic Models. https://papers.nips.cc/paper/2011/file/5ef698cd9fe650923ea331c15af3b160-Paper.pdf
13. Akiva, N., Greitzer, E., Krichman, Y., & Schler, J. (2021). Mining and Visualizing Online Web Content Using BAM: Brand Association Map. Proceedings of the International AAAI Conference on Web and Social Media, 2(1), 170-171. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/18635
14. S. Erevelles, N. Fukawa and L. Swayne Journal of Business Research, 69 (2) (2016), pp. 897-904
15. C.D. Manning and H. Schütze MIT Press, London, UK (1999) Foundations of statistical natural language processing

16. W.S. El-Kassas, C.R. Salama, A.A. Rafea, H.K. Mohamed, Automatic text summarization: A comprehensive survey, Expert Syst. Appl. 165 (2021) 113679, http://dx.doi.org/10.1016/J.ESWA.2020.113679.
17. He, X., Zemel, R. S., & Carreira-Perpinan, M. A. (2004). Multiscale conditional random fields for image labeling. In Proceedings IEEE conference computer vision and pattern recognition.
18. Ladicky, L., Sturgess, P., Alahari, K., Russell, C., & Torr, P. H. S. (2010). What, where and how many? Combining object detectors and CRFs. In Proceedings European conference computer vision (pp. 424–437).
19. Shotton, J., Johnson, M., & Cipolla, R. (2008). Semantic texton forests for image categorization and segmentation. In Proceedings IEEE conference computer vision and pattern recognition.
20. Galleguillos, C., & Belongie, S. (2010). Context based object categorization: a critical survey. Computer Vision and Image Understanding, 114(6), 712–722.