

MSc Data Science and AI for the Creative Industries

Natural Language Processing for the Creative Industries

Unit Code: IU000131

Element 1: Critical Essay

Building a Fashion Outfit Generator

Nadia Alsagoff (ALS16494646)

Date: 22 Jan 2021

Word Count: 2050

<b>Background</b>	<b>2</b>
Fashion Forecasting	2
Clothing Recommendation	3
Concerns with Research on Fashion-Tech	3
<b>Project Aims</b>	<b>4</b>
<b>Methods</b>	<b>4</b>
Areas of Concern	5
Fashion Terminology	5
N-grams	6
Distinguishing Categories	6
<b>Results</b>	<b>6</b>
Colour	6
Material	7
Pattern	8
Style	9
<b>Discussion</b>	<b>10</b>
<b>References</b>	<b>12</b>

## **Background**

### **Fashion Forecasting**

Common industry tools for fashion forecasting include WGSN and Edited which are paid subscription services.

WGSN takes a creative driven approach with trend boards developed by experts in the industry. Their global team produces over 250 new trend reports each month, planning colours and styles up to 2 years in advance (WGSN, 2021). Experts at WGSN also consider the impact of the macro-environment on fashion such as politics, technology and culture in society as a whole.

While WGSN's content is manually curated, Edited uses AI to scrape online marketplaces and track product SKUs. Their engine, AtlasAI, "uses visual pattern recognition, optical character recognition and text crawling to collect data" (Edited, 2021). AtlasAI is able to provide information on category, colour, price, pattern etc with industry-standard terminology. With over 8 years of data stored, companies can see overviews of trends overtime, in different markets and across competitor brands.

In Silva et al. (2019), Google Trends was identified as a key source for analytics and fashion forecasting. It's dataset is gathered from user web searches to find out what customers are looking for. This is more of a bubble up form of forecasting as it anticipates what will be in demand.

The paper was critical of WGSN, suggesting that with more companies subscribing to the service, the designs appearing throughout fashion brands are beginning to look similar due to inspiration coming from the same source. It also discusses how relying on Edited leads companies to focus too much on competitors over customer needs.

Fashion forecasting has always been a challenge due to the volatile nature of consumer demand. The TrendFashion tool for identifying fashion trends was developed in Beheshti-Kashi et al. (2015). The system "identifies fashion related words and weights them according to an index". It crawls German fashion websites before processing the data through a tokenizer, named entity recognition and POS tagger. The paper concludes by noting that even slight differences in colour or silhouette can make a large impact on the overall look and success of a product. This leads to the rationale that elaborate product description and details are crucial in fashion datasets.

## Clothing Recommendation

Many current clothing recommendation softwares use images for their learning data, using CNN to extract low-level features like category and length. In the paper by Yu et al. (2018), this is taken a step further to include aesthetics. It emphasises the lack of research in extracting high-level aesthetic features plus the value it has in improving clothing recommendation systems. Using a tensor factorisation model, they were also able to account for user preference based on past history as well as time. This meant that if a user purchased a dress, the model should recommend a similar dress in reference to the style of that previous purchase while also being relevant to current fashion trends.

In Lin et al. (2019), a deep learning-based framework (NOR), was proposed to give outfit recommendations. They used CNNs to model tops and bottoms as latent vectors before extracting visual features by employing the top vectors to match the bottom. The result is a rating score of the match. Then, an RNN transformed these features into a sentence. Content from fashion website Polyvore was used as the dataset to test this framework where they found that sentence generation never included any 'mismatch' as the comments on Polyvore were rarely negative. (Lin et al., 2019)

## Concerns with Research on Fashion-Tech

TABLE 9: Examples of recommendations and generated comments.

wow ! this is so beautiful ! love the skirt ! (✓)	love the pink ! (✓)	great denim look . (✓)	love the color combination ! (✓)
love this set ! the colours are amazing . (✓)	so beautiful and such a nice style here like it . (✓)	great look great set great color . (✓)	love the red and white ! (✓)
great look great set great mixing outfits n ' nice bag . (✗)	thank you so much for your lovely comments ! (✗)	congrats on top sets sweetie ! xxo . (✗)	great set , love the shoes ! (✗)

Despite these advancements in research, it would still be appropriate to question the success of these experiments given that there rarely are creative experts to weigh in on the results. As seen in the output above, the model was trained on a Polyvore dataset, a website that ceased to exist since 2018 and is supplied by user generated content instead of fashion experts. The lack of diversity in computing is also brought into question specifically in the research of womenswear fashion recommendations.

The decline in percentage of women in IT is unsurprising due to obstacles faced in entering the industry. This includes issues in the social culture and “pandemic objectification of women” in the tech industry (Branson, 2020). A study proved that although there was no significant difference in productivity between male and female developers, there were still technical biases against female developers with lower code acceptance rates. The groups studied also had less than 10% female developers in each project (Bosu and Sultana, 2019).

## Project Aims

**Aim:** Create a model that will plan and style an outfit based on the items in a user's closet.

The overproduction of fashion has increasingly become a major environmental issue, with fast-fashion being the main cause for this excess waste. The speed of product turnover together with the decreasing quality of garments has led to an industry that favours profit first. Fashion designer Phoebe English describes fast fashion as a “monstrous disposable industry”, with global apparel consumption projected to rise by 63% by 2030 to 102 million tons (House of Commons UK, 2019).

Unfortunately, this oversupply by companies is moulding a habit of throwaway fashion in consumers in a cycle that drives prices lower and demand higher (i.e quantity over quality). According to the journal Environmental Health, consumers need to adopt a ‘less is more’ mentality. In 2018 it was reported that nearly 85% of clothing bought in America is sent to landfill every year, making up almost 80 pounds per person (Bick, Halsey and Ekenga, 2018).

This project seeks to create a tool that helps reduce throwaway culture by helping users rediscover their wardrobe and encouraging ‘Reducing’ purchases and ‘Reusing’ existing items. This is in contrast to the multitude of projects funded that aim to do the opposite by encouraging new purchases and pushing sales of new items.

## Methods

**How:** Transcribe all the articles of clothing in a wardrobe and input them as ‘rules’ in Tracery. Create a classifier to recognise style.

This paper presents a basic framework for mixing and matching a user’s wardrobe followed by recommending the appropriate style to wear the proposed outfit.

## **Areas of Concern**

In the journal by Vaccaro et al. (2018), the focus was on improving a fashion-focussed chatbot by running a wizard-of-oz experiment. Some of the needs are listed below.

1. Advice on dressing for specific occasions. (eg, conference, wedding, winter)
2. Help in reconciling their personal style with their aspirations.
3. Matching or pairing a specific item. (eg. 'I have this, how do I wear it?').
4. The least prioritised factor in personal styling was looking 'trendy'. Clients more often than not wanted clothes for the everyday instead of following current trends.

Stylists would reply to users with images and text descriptions to show ideas. It was found that images alone was not enough as it could not accurately communicate details such as fabric. There was also a positive to having a 'third-party' stylist like a chatbot as personal in-store or commission-based stylists would be biased in their product offer to make sales.

## **Fashion Terminology**

The importance of detailed fashion terminology in NLP is also proved in Kotouza et al. (2020), where they developed a system to act as a personal assistant for fashion designers. It first uses NLP to analyse information extracted from images of clothing, then passes this through a model to suggest clothing designs. The dictionary was created by experienced fashion designers and lists all possible accepted values for a specific attribute. Words are labelled by matching the meta data to the dictionary. (Kotouza et al., 2020)

Having a standard set of expert terminology helps to keep vocab accurate and efficient. (eg. names of skirt: wrap/pleat/tulip/a-line/circle etc). This also takes into account the variance by country (eg, 'romper' in US vs 'playsuit' in UK).

This is relevant in the 'set-up' stage of transcribing the wardrobe, as well as in the fashion dataset. It is questionable how user-friendly this is for an average person to transcribe their wardrobe themselves without an expert.

Another example of this would be words with double meanings. The tokenizer used in the code did not lemmatise or remove common words as this might change meaning entirely (eg 'short' to describe length vs 'shorts' as in an item of clothing). The dataset made sure to use words like 'midi', 'mini' and 'crop' to describe length.

## N-grams

To some extent, the more descriptive a garment is written out, the easier it is to identify and categorise. This does have its limits, most likely representing a bell shaped curve. At a certain point, too many words makes the garment too ‘unique’. It is useful to have some consistency in the number of words used to describe a garment, in line with the number of n-grams chosen.

With the first trial of the model, it achieved a high accuracy rating of over 90%. However, this was a false positive as it was revealed that high frequency words were not necessarily useful in identifying style. This included words like ‘long’ and ‘sleeve’ being separated as well as ‘high’ and ‘heel’. Descriptors like these are much more beneficial when joined together than on their own.

## Distinguishing Categories

For this project, 3 different styles were labelled. Formal/Casual/Party. Despite the accuracy rating falling to just under 90%, this was accepted as style is fluid and can easily overlap. Mentioning colour in the dataset was also carefully considered. Minimalist colours like ‘black’ and ‘white’ are more often in the ‘formal’ category as they are more likely to fit formal situations compared to ‘yellow’.

However, without context, it would be inaccurate to use colours as absolutes because there will likely be exceptions. Instead, features like silhouette and material would be more useful. I.e ‘glitter’ for Party or ‘oversized’ for Casual.

## Results

### Colour

Since using Tracery meant that the machine was not learning language, the grammar needed to be carefully built. In this project, clothes were separated by 'Light' and 'Dark' with the theory that keeping colour shades together would increase chances of a match.



Start with your maroon knit sweater. Pair it with a darker colour like your black denim crop jeans. For shoes, wear your black highcut converse. Layer with your black logo sweater. -> Style: Casual

The limitation of this is that a simple 'white t-shirt with black jeans' would never be a possibility. To expand this, a more elaborate form of dividing the wardrobe by category would be needed as well as more sentence generators.



Start with your orange high neck pleated blouse. Pair it with a darker colour like your red bodycon midi skirt. For shoes, wear your mustard low heels. Layer with your green corduroy sweater. -> Style: Formal

## Material

This method also disregards fabric and texture. In the example below, a lightweight sheer top was paired with a bulky corduroy trouser. Although the colours matched, there was a disregard to proportions in material.



Start with your white sheer short sleeve shirt. Pair it with a lighter colour like your sage corduroy crop trouser. For shoes, wear your black strappy ribbon heels. Layer with your light blue chunky knit sweater. -> Style: Casual

However, this would be useful for creative purposes in idea generation as Tracery would provide outputs that push the boundaries and could sometimes work.



Start with your sage vneck cami. Pair it with a lighter colour like your white denim biker short. For shoes, wear your navy suede docs. Layer with your white denim jacket. -> Style: Casual

## Pattern

Controlling the mix of patterns is complicated as this varies greatly depending on the visual appearance.



Start with your maroon paisley short-sleeve shirt. Pair it with a darker colour like your black chevr on mini skirt. For shoes, wear your black stiletto heels. Layer with your navy sweater. -> Style: Par ty

While this could be remedied by confining the grammar to pair a print with a plain, there could also be an opportunity cost in doing so.



Start with your black tshirt. Pair it with a darker colour like your green floral slit midi skirt. Fo r shoes, wear your black strappy floral heels. Layer with your black logo sweater. -> Style: Formal

## Style

As mentioned, there could be overlaps in style. Although the trousers below are technically formalwear, the model managed to place this as 'Party'.



Start with your black bralette. Pair it with a darker colour like your black crop trouser. For shoes, wear your mustard low heels. Layer with your navy corduroy oversized shirt. -> Style: Party

It is also important to consider the weights of certain features. It was found that anything that included high heels tended to 'overtake' the probability ratio and place outfits in 'Party' despite the rest of the outfit. Although playing with this in the dataset can lead to favourable results.



## **Discussion**

Fashion has a huge sustainability issue. Several companies who have tried to address this have even been accused of greenwashing in a bid to actually sell even more clothes. The conversation has been too focussed on recycling and too little on the first two R's, 'Reduce' and 'Reuse'. With this framework, the project succeeds in preventing bias in regards to advertising and brands pushing for sales. This offline method of recommendation is completely user oriented and disregards fast-fashion tactics.

However, in its current state, negative matches are still common. While some abstract pairings could work, not all of them are positive matches. This has also been tested on one wardrobe and it cannot be determined whether this will work as well with different inputs.

In the future, this code above can be built upon to include more columns and categories for the Tracery grammar. Outputs can be added back into the dataset to expand it and more labels can be included such as splitting casual into 'formal casual' and 'leisure'.

## References

- Beheshti-Kashi, S., Lütjen, M., Stoever, L. and Thoben, K., 2015. *Trendfashion - A Framework For The Identification Of Fashion Trends*. Bonn: Gesellschaft für Informatik, pp.1195-1205.
- Bick, R., Halsey, E. and Ekenga, C., 2018. The global environmental injustice of fast fashion. *Environmental Health*, [online] 17(1). Available at: <<https://ehjournal.biomedcentral.com/articles/10.1186/s12940-018-0433-7>> [Accessed 3 January 2021].
- Bosu, A. and Sultana, K.Z. (2019). *Diversity and Inclusion in Open Source Software (OSS) Projects: Where Do We Stand?* 2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM). [online] Available at: <https://ieeexplore.ieee.org/abstract/document/8870179> [Accessed 15 Jan. 2021].
- Branson, D., 2020. *The Future Of Tech Is Female: How To Achieve Gender Diversity*. New York: NYU Press.
- EDITED, 2021. *EDITED Data | The Market Intelligence Suite | Global Retail Data*. [online] EDITED. Available at: <<https://edited.com/products/market-intelligence-suite/edited-data/>> [Accessed 15 January 2021].
- Gallagher, C., 2019. *Obsessive Consumption Disorder*. [online] Fashion Revolution. Available at: <<https://www.fashionrevolution.org/obsessive-consumption-disorder/>> [Accessed 1 January 2021].
- Kotouza, M., Tsarouchis, S., Kyprianidis, A., Chrysopoulos, A. and Mitkas, P., 2020. Towards Fashion Recommendation: An AI System for Clothing Data Retrieval and Analysis. *IFIP Advances in Information and Communication Technology*, [online] pp.433-444. Available at: <[https://link.springer.com/chapter/10.1007/978-3-030-49186-4\\_36](https://link.springer.com/chapter/10.1007/978-3-030-49186-4_36)> [Accessed 3 January 2021].
- Lin, Y., Ren, P., Chen, Z., Ren, Z., Ma, J. and de Rijke, M., 2019. Explainable Outfit Recommendation with Joint Outfit Matching and Comment Generation. *IEEE Transactions on Knowledge and Data Engineering*, [online] 32(8), pp.1502-1516. Available at: <[https://www.researchgate.net/publication/331882619\\_Explainable\\_Outfit\\_Recommenda](https://www.researchgate.net/publication/331882619_Explainable_Outfit_Recommenda)n\_with\_Joint\_Outfit\_Matching\_and\_Comment\_Generation> [Accessed 2 January 2021].

Parliament.uk. 2019. *Fixing Fashion: Clothing Consumption And Sustainability - Environmental Audit Committee - House Of Commons*. [online] Available at: <<https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1952/report-files/195204.htm>> [Accessed 1 January 2021].

Silva, E., Hassani, H., Madsen, D. and Gee, L. (2019). *Googling Fashion: Forecasting Fashion Consumer Behaviour Using Google Trends*. Social Sciences, 8(4), p.111

Vaccaro, K., Agarwalla, T., Shivakumar, S. and Kumar, R., 2018. Designing the Future of Personal Fashion. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, [online] pp.1-11. Available at: <[https://s3.amazonaws.com/kvaccaro.com/documents/CHI18\\_fashion.pdf](https://s3.amazonaws.com/kvaccaro.com/documents/CHI18_fashion.pdf)> [Accessed 16 January 2021].

WGSN. 2021. *Key Trends: 2021 And Beyond*. [online] Available at: <<https://www.wgsn.com/en/>> [Accessed 15 January 2021].

Yu, W., Zhang, H., He, X., Chen, X., Xiong, L. and Qin, Z., 2018. Aesthetic-based Clothing Recommendation. *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, [online] Available at: <<http://dx.doi.org/10.1145/3178876.3186146>> [Accessed 1 January 2021].

#### **Github:**

Allison Parrish, *Tracery for Python*, Github Repository, <https://github.com/aparrish/pytracery>

Kate Compton, *Tracery*, Github Repository, <https://github.com/galaxykate/tracery>

Ole Schütt, *A Jupyter extensions that turns notebooks into web applications*, GitHub Repository, <https://github.com/oschuett/appmode>