



Zara Smart Naming

Camila Barbagallo, Ryan Daher, Rocío González Lantero, Joaquin Maroto

IE University

Professor: Noa Cruz Díaz

<b>Abstract</b>	<b>3</b>
<b>Key Information</b>	<b>3</b>
<b>Introduction</b>	<b>4</b>
<b>Related Work</b>	<b>5</b>
<b>Experiments</b>	<b>6</b>
Data	6
Evaluation Method	6
Experimental Details	6
<b>Results</b>	<b>7</b>
<b>Analysis</b>	<b>8</b>
<b>Conclusion</b>	<b>8</b>
Limitations	8
Discussion	9
<b>Resources</b>	<b>10</b>
<b>Appendix</b>	<b>11</b>
Appendix 1	11
Appendix 2	11

## **Abstract**

The world of automated fast-fashion has been rapidly growing in size, inviting with it the applications of data science and machine learning. This study is oriented around the problem of product name generation using a product description proposed by Zara through a hackathon competition where participant results are to be used as benchmarks. This research focused on Google's GPT-2 algorithm and its re-training to suit the competition and data requirements. The final results include a re-trained GPT-2 model capable of learning from new data, and able to generate product names from their description. Future works are centered around increasing resources to operate at full potential and attending new competitions to compare results.

## **Key Information**

[Here](#) you can find the code we used for this project. (Google)

## Introduction

NLP, also known as Natural Language Processing was first introduced in 1950, aiming to study the capability of a computer to mimic human intelligence under certain conditions. Within NLP there exist two main professions: computational linguistics and NLP engineers. Computational linguistics refers to people who have studied philosophy and NLP engineers to those who have studied computer science. The main benefits of NLP are that it allows the processing and analysis of multiple text data and the structuring of a highly unstructured data source. On the other hand, some of the problems NLP faces are inconsistencies, grammatical errors, and morphological, syntax and semantic ambiguity. 90% of the available data is in textual format, making NLP an essential asset for any company in any industry. With the use of NLP, the fashion industry could improve the customer experience, offering its clients, for example, a personal AI stylist, which would give opinions on outfits based on trends and professional knowledge (Pupillo, 2019). Besides, just as companies in other industries, fashion companies could optimize their processes with the use of NLP, accelerating the delivery of their products.

Moreover, natural language processing involves two different processes: data preprocessing and algorithm development. On one hand, data preprocessing deals with cleaning and preparing the text data to place data in workable form and highlight features to allow algorithms to analyse it. There are three main features dealing with data preprocessing. To begin with, tokenization is used to break down text into smaller units. Moreover, stop word removal eliminates common words, such as “a”, “the”, “and”, etc... and only selects the most identifiable and informative words. In addition, lemmatization and stemming is another preprocessing method, used to reduce words to their root forms. Finally, part-of-speech tagging tags words based on their type (noun, adjectives, adverbs...). On the other hand, once data has correctly been processed, algorithms are used to analyse it. There are two commonly used algorithm types (for more algorithms check *Method* section): rule-based and machine learning-based. Rule-based algorithms are used in linguistic rules and are the simplest form of analysing NLP. It uses prescribed knowledge-based rules to solve a problem. Machine learning-based algorithms are used to analyse human language and use statistical methods to learn and perform based on pre-trained data. There are many other algorithms and in reality, using a combination of deep learning, machine learning, neural networks and natural language processing algorithms is the best option to obtain the most accurate results.

In addition to the direct benefits that AI, and NLP can have on the fashion industry, the act of identifying tags, names, and descriptions can serve tremendous benefits to numerous industries. These applications save a large amount of time that would be spent on manual tagging, encouraging more streamlined and optimized procedures, and facilitating the searching of tiles and products. Its various applications include product tagging for easier searching, item/person tagging for clustering and type identification, as well as website tagging for optimized SEO ranking. By sending images, or descriptions through a machine learning pipeline, a set of ready-to-use tags can be returned from the system with near-immediate business impacts. Although the benefits are numerous, it is also important to ensure a properly built information infrastructure. As explained by PicturePark, an AI-based Content Manager, the current state of AI systems seem to perform much weaker than expected as most algorithms in use today have trouble differentiating between crucial

descriptive or image components, and other appearing ones as “they lack historical or advanced situational context” (Forster, 2018). It is also important to ensure that the produced tags are ones that the company wants to be associated with, “If for instance, you are a petrochemical company and your refinery pictures at sunset are all tagged with ‘pollution’ then you might want to consider if publishing these pictures are a good thing.” (Forster, 2018).

Zara is considered “the leader of its market”, relying on two success factors: technological innovation and customer experience. Their main strategy is to deliver new collections every two weeks whereas other brands only market 6 collections a year. Through technology, they are not only able to improve their customer experience at many different levels, as controlling the temperature in all its shops worldwide from one same place, but it also enables them to optimise their operations by, for example, predicting which products will be easier to sell and the quantities they need to ship to each shop. Like Zara, the fashion industry has a lot of data about their customers and slowly more companies are following Zara in using it (IE Insights, 2020). Implementing NLP techniques could help Zara speed up the creation of the new collection, by automatically giving each product a name from its description

The main objective of this project is to help Zara keep on optimizing its process by automatically tagging new items from their description. This will speed up the commercialization of new collections, allowing Zara to have more time for transport, design and confection. Furthermore, due to the nature of this being a competition proposed by Spain AI in collaboration with Zara, we aim to reach the top 10th percentile before the end of April.

The goal will be accomplished by developing an NLP model capable of generating short names based on Zara’s clothes’ descriptions. The model will learn Zara’s way of naming its products based on their previous collections. Moreover, this task has been set as a competition by Spain AI in collaboration with Zara Tech under the name “Reto NLP de Spain AI” (Álvaro, 2020). Results will be evaluated by a complete match and using DCG. Participants can submit a maximum of 5 times per day and 100 throughout the whole competition. Currently, the maximum score is 37.85 DCG, which demonstrates the complexity of the task.

The difficulty behind this project is that there are no previous studies regarding product name creation. Therefore we need to use existing text generation and text summarization algorithms and tune them accordingly so we can get the result we wish.

## **Related Work**

It is important to note that research has been done regarding similar tasks but nothing was found. The models used are quite new, so they do not have a lot of clear documentation. Therefore, we have used the competition’s website as guidance and benchmark to compare the results with the ones of other participants.

## Experiments

### Data

From the competition, we have received two files, the training set and the test set, which are both in English. The training file (Appendix 1) includes 33,613 observations, with two columns, one with the product name and the other with the description of it. On the other hand, the test dataset (Appendix 2) includes 1441 descriptions from which we need to predict the name. With the data available, the product descriptions will be the input and we will get the product names as output.

### Evaluation Method

To evaluate our models we will be using the Discounted Cumulative Gain (DCG), which is the metric that has previously been established by the competition. DCG is usually used for information retrieval algorithms, measuring the similarity of the query to the information given. In this case, we will be measuring the similarity of our predicted name to the actual name, and our goal will be to maximize it (S., 2017). Moreover, we will be also measuring the accuracy of the model, by only validating as a correct prediction when the model predicts the exact same product name as the actual one. To ensure valid results, we will split our training dataset into training, testing, and validation in order to have lengthy training, fine-tuning using the test set, and final metric outputs using the validation.

### Experimental Details

After taking the first look at the data, several factors stood out to us. The main ones being all of the punctuation, numbers and possible stop words. The first thing we did was to delete the HTML tags, as we noticed some can be seen in the example of the rows within the test dataset. The next thing we did was to delete unnecessary white spaces, although from a first glance we didn't notice any. However, we felt it was a step that could be taken and if there weren't any white spaces nothing would change. Following these two steps, we took away the accents and converted all characters to lowercase. All of this was done after tokenizing the text, then we converted these tokens to their base form and deleted all words belonging to our stopword list, which was the default one for spacy in English.

Among our first set of models, we focused on Recurrent Neural Networks, a fully connected deviant of feedforward neural networks. What is special about RNNs is their simplified use of stored memory within the model. They perform a function for every data that is inputted which is dependent on the previous inputted cell. This process is repeated and propagated forward through the data as the produced output is then inputted back into the network to impact the transformation of the following output. Finally, it bases its decision on the inputs provided and their relation to one another as part of a sequence of a text.

Moving onto the next technique, we have the LSTM network which serves as a special type of RNN. They seek to make it easier to remember past data stored in memory. When given time lags of unknown durations or gaps, LSTMs are well equipped to process, classify, and predict different sequential based information such as the description text provided to us. It processes the information using a series of filters: the input gate, drop/forget gate, and the output gate. Through them, they are able to circumvent the main issue of vanishing gradients while generating the output target. The main hurdles however of

LSTMs are the fact that a complex path links the previous cell to the current one, limiting the amount of information that can be passed through it and remembered. The second is that the higher complexity of these models, result in higher computational requirements; and finally, the sequential nature of LSTMs lead to harder parallelization, again increasing processing times, particularly with larger datasets.

The aforementioned traditional RNN based methods would be compared to more dynamic and unique methods such as the Attention Mechanism. The paper “Attention is All You Need” which was researched and published by Google Research showcases newer attention based techniques that allow higher parallelism and thus more efficient and faster results. Through encoding and decoding mechanism, the transformer is able to encode relevant word context into a word vector allowing the network to “zoom in and focus on relevant contextual words in both the input sequence and the outputs predicted up to that point, to determine the next output” (Zappy, 2020). Although one of its main uses is in language translation, it may prove to be beneficial to our scenario, providing particular words more attention within their context and allowing for more precise outputs that could link to Zara’s word bank.

Finally, for close to state-of-the-art methods we worked with the GPT-2 (Generative Pre-trained Transformer 2) model. This is a text-generation model developed by OpenAI in 2019. GPT-2 is built out of many layers of decoders, which finally output one token at a time, which is then added to the sequence of inputs, similar to the RNNs mentioned above. Each decoder is composed of a feed-forward neural network and a masked self-attention layer. The self-attention layer makes sure the model is understanding the words that are referring to one another by assigning relevance scores to each word and adding up the vector representation. There are three key components for this layer, the query, which is the representation of the current word we are scoring against the rest, keys, which are the labels for all the words in the segment and the value, which is the value given to each word by the layer. After this, the outputs of this layer become the inputs of the feed-forward neural network, which later generates the text. The word generated is selected by calculating the probability for the model to choose each word. When training, GPT-2 uses start and end tokens you have to define. For the challenge at hand, we also needed to add an extra token, which split the description from the product name, indicating the model it needed to predict that part (Alammar). For this model we were able to run one single epoch, due to our computing resources.

## **Results**

The objective of this project was to use and experiment with state of the art NLP models in an attempt to predict clothing names based on their description. Among the various methods that we have attempted, GPT-2 seemed to perform the best and have more robust and dynamic prediction capabilities. Although it is one of the most advanced and best performing natural language processing models, it requires intensive resources to which we did not have access. With this being said, we were able to preprocess our data to fit the functionalities of the GPT model by merging and modifying the text input to include special strings identifiable by the model. Through this, we were only able to train our model using a single epoch. The training time took two and a half hours for the first epoch and gave us a loss of 0.06 in our test set. Within our left-out validation set of 3352 products our model was able to obtain an accuracy of around 5%, indicating that the model was able to fully predict

5% of the product names given their description. Although this figure may be low, especially when compared to some of the top competitors, we believe that with more resources and more training time, our model would be able to learn very quickly and perform much better than at its current state.

## **Analysis**

The GPT-2 model was trained with over 40GB of internet text, and has over 1.5 billion parameters. While this robustness and high number of pre-trained data vastly increases the performance of this NLP model, it also requires a tremendous amount of resources for length training. The original task of this model was to predict a word given all the previous words within a text, and this simplicity has allowed the model to be dynamically used in different scenarios. This has allowed this model to be more easily reconfigured to suit the need of the Zara problem through GPT's encoder and special tokens. However, within stricter constraints it could be beneficial to combine the text generator to a text summarizer algorithm to obtain results more suited to the particular use case. Furthermore, we think more research should be done in terms of the inner workings of the model, so the documentation can be improved. The goal would be to create a model, which is not a classification model, but to which you can show to create outputs with a certain length and it learns from the outputs.

## **Conclusion**

Through this project, we have been able to research and to experiment with various NLP algorithms, including state-of-the-art models such as OpenAI's GPT-2. It has been interesting to dive deeper into the architecture of such a dynamic yet robust algorithm created by some of the world's best engineers and scientists. We have been able to use the GPT-2 encoder to transform our dataset into a parameter set, establish special tokens to indicate particular queues, train the model on Zara's dataset, and generate product names for future product descriptions. However, the model itself is very large and requires substantial computing capabilities, and so we have not been able to test out the model's performance at its best potential. In addition, Zara needs a model with more stringent constraints, able to learn their own technique of naming their product. This would take into account own Zara words, such as TRF and also the length of the product names.

## **Limitations**

During this project, we have faced several different challenges. The main issue we faced in this challenge was not having enough documentation with regards to the model or the theory. With this in mind, it was extremely difficult to solve all the issues faced throughout the project with the code. We believe that if more time would have been given to this task we would have been able to research more on the GPT-2 model and would have been able to improve our scores on this challenge. We also weren't aware of the change in the competition of having a due date so we were unable to submit our model through the website. This didn't allow us to complete the project to the end. Finally, due to the lack of computing resources, we were limited to only one epoch of training, which itself took about two and a half hours. With more resources, our model would have been able to train for a few epochs and retrieve much better results.



## Discussion

The fashion industry is undergoing a huge transformation, moving from a creative and advertising realm, to one of technology and data processing. With companies like Zara looking to take advantage of current digitalization, new advancements to the field of data science within fashion has begun to take off, and create entirely new business models and clothing lines. With the coinciding rise of social networks, there has never been access to a larger pool of information coming directly from consumers. This blend creates the perfect opportunity for NLP to dominate in the fashion realm, capturing and understanding customer information and opinions and transforming them into tangible business models. Within our project, we have been able to use some of the most advanced pre-made models to create product names using only their description. While advanced and top-performing, these models require large amounts of resources and so we have only been able to test a tiny fraction of their potential. By using a cloud platform such as AWS, or Azure, we will be able to further train the GPT model over multiple epochs and identify their true predictive powers within the Zara naming context.

## Resources

Alammar, J. (n.d.). The illustrated gpt-2 (visualizing transformer language models). Retrieved April 16, 2021, from <https://jalammar.github.io/illustrated-gpt2/>

Álvaro, N. (2020). Demuestra Tus HABILIDADES en IA. SPAIN AI. Retrieved from [http://www.spain-ai.com/hackathon2020\\_reto\\_NLP.html](http://www.spain-ai.com/hackathon2020_reto_NLP.html)

Forster, R. (2018, July 18). Artificial Intelligence in Content Management: Fact, Fiction and Future. Retrieved March 02, 2021, from <https://picturepark.com/content-management-blog/best-practices-for-dam-ai-auto-tagging-content-management-blog>

Google. (n.d.). GPT-2 Fine-Tuning Tutorial with PyTorch & Huggingface in Colab. Google. <https://colab.research.google.com/drive/13dZVYEOMhXhkXWfvSMVM1TTtUDrT6Aeh?usp=sharing#scrollTo=EKOTlwcmxmej>.

IE Insights (2020, March 12). Zara: Technology and user experience as drivers of business. Retrieved February 28, 2021, from <https://www.ie.edu/insights/articles/zara-technology-and-user-experience-as-drivers-of-business/>

Pupillo, M. (2019). Artificial Intelligence and the Fashion Industry. LUISS Department of Economics and Finance. Retrieved from [http://tesi.luiss.it/25378/1/212661\\_PUPILLO\\_MARIANNA.pdf](http://tesi.luiss.it/25378/1/212661_PUPILLO_MARIANNA.pdf)

S, S. (2017, July 24). Discounted cumulative gain. Retrieved March 08, 2021, from <https://machinelearningmedium.com/2017/07/24/discounted-cumulative-gain/>

Zappy, T. (2020, June 07). The current state of the art in natural language processing (nlp). Retrieved March 10, 2021, from <https://zappy.ai/ai-blogs/the-current-state-of-the-art-in-natural-language-processing-nlp>

## Appendix

### Appendix 1

	name	description
0	CROPPED JACKET TRF	Jacket made of a technical fabric with texture...
1	OVERSIZED SHIRT WITH POCKET TRF	Oversized long sleeve shirt with a round colla...
2	TECHNICAL TROUSERS TRF	High-waist trousers with a matching elastic wa...
3	SHIRT DRESS	Collared dress featuring sleeves falling below...
4	PUFF SLEEVE DRESS WITH PLEATS TRF	Loose-fitting midi dress with a round neckline...

### Appendix 2

	description
0	Knit midi dress with a V-neckline, straps and ...
1	Loose-fitting dress with a round neckline, lon...
2	Nautical cap with peak.  This item mus...
3	Nautical cap with peak. Adjustable inner strap...
4	Nautical cap with side button detail.  ...