Methodology Used

**Transfer Learning**

Humans can learn tasks and use the skills gained, thus, for some other jobs. The more similar the given assignment is to the prior one, the more comfortable one is in its execution. Conventional ML and deep learning algorithms, until a few years back, could only work in isolation. Every time a new dataset presented itself, the model had to be developed from scratch. Transfer learning is the learning paradigm to overcome the isolation, as mentioned earlier, and utilize the knowledge previously acquired for one task to solve similar ones. Granted Transfer Learning is a fantastic idea to improve deep learning models, but it has its limited uses. It has been working great in the area of Image classification, but when it came to the domain of Natural Language Processing, it couldn't help much. 2018 though, was the exciting moment in the field NLP. As Sebastian Ruder puts it, "NLP's ImageNet moment has arrived," and it sure has. Methods such as ULMFiT, OpenAI GPT, ELMo, and Google AI's BERT have a significant impact on transfer learning in NLP by using language modeling during pre-training. The interesting aspect of the transfer learning methods mentioned above is that they use language models pre-trained on well-formed, massive curated datasets that include full sentences with a clear syntax (such as Wikipedia articles and the 1 billion word benchmark). The given kernel implements the ULMFiT model.

Features Extracted

* review\_title
* review\_description
* country
* province
* region\_1
* generated features such as ‘value for money’[points/price]

Models Accuracy and Losses

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models | Training Accuracy | Training Loss | Validation Accuracy | Validation Loss |
| ClassifierModel1 | 76.7000 | 0.72535 | 76.6691 | 0.725286 |
| ClassifierModel2 | 97.7566 | 0.0639339 | 97.3987 | 0.0713948 |
| ClassifierModel3 | 97.7876 | 0.0624435 | 97.5318 | 0.0729472 |
| ClassifierModel4 | 97.8333 | 0.0605756 | 97.7374 | 0.0657355 |
| ClassifierModel5 | 97.8951 | 0.0585178 | 97.8221 | 0.0629165 |
| ClassifierModel6 | 97.7163 | 0.0621578 | 97.5681 | 0.0713805 |

Though it’s difficult to differentiate among Model 2 to 6, Model 5 generated the best accuracy

Exploratory Data Analysis

1. Which Countries Produce the wines rated highly by sommeliers?

A picture containing man

Description automatically generated

India features in top 5 wine producing countries, whose wines are lauded by sommeliers

1. Which Countries Produce the Costliest wines?

A picture containing sitting, black

Description automatically generated

* Europe leads the charge in producing the costliest wines.
* Switzerland is no surprise as it is one of the countries where standard of living is pretty high. This can be attested from the fact that the top 2 countries(Swiss and English) have a huge difference between them

1. Which countries produce Value for Money wines?

A picture containing black, man, water, white

Description automatically generated

* None of the traditional wine producing countries(France, US, Italy) features in top in this list
* India once again feature in top 5

1. Which Winery in which Country is the most costliest?

A close up of a logo

Description automatically generated

* France and US, the traditional wine producing countries, as expected make in to this list

1. What words describe the top 3 highly rated wines?

A black sign with white text

Description automatically generated

* Nebbiolo seems to be associated more with the word pepper and spice. I wonder what would it taste like
* Grüner Veltliner and Champagne Blend are associated with acidity and citrus. It's may be because their taste is similar to lemons