Personality Detection of Movie Characters based on Movie Scripts

Team: Saurabh Raut, Jay Gala, Shivani Reddy Rapole.
Project Mentor TA: Pratik Kunapuli

1) Abstract

In this project we are trying to understand the personality traits of movie characters based on movie scripts. Our primary contribution is training various machine learning models that predict the Big Five personality traits and applying them on movie scripts to understand the character's emotions in each scene. We trained and tested a few linear and non-linear models on text data that predicts the Big Five traits (OCEAN model) - openness, conscientiousness, extraversion, agreeableness, and neuroticism. After evaluation we found that a very simple MLP (multi-layer perceptron) model with BERT embeddings generated good results and hence applied the same on scene segmented movie script dataset to analyse scene wise personality of each character.

2) Introduction

The same actor/actress might be portraying different personalities in different movies. Moreover the character's personality may also vary across different scenes in the same movie. Reading the entire movie script to understand the character's personality might be too tedious. To make this process easier, in this project, we perform a scene-wise personality trait prediction depending on the dialogues of the character in that particular scene in the form of OCEAN model predictions. Our primary contribution is collecting datasets from a different domain, comparing the performance of various ML models including linear and non-linear models, and finetuning it on a movie's dataset. So, the input to each of our five binary classifiers (for each trait) is the set of dialogues of a character in a scene and the output is presence of the corresponding Big Five traits (OCEAN model) - openness, conscientiousness, extraversion, agreeableness, and neuroticism.

1. Extroversion (EXT) - Is the person outgoing and energetic versus reserved?
2. Neuroticism (NEU) - Is the person sensitive and nervous versus confident?
3. Agreeableness (AGR) - Is the person trustworthy, straightforward, generous, and modest versus unreliable, complicated, meager, and boastful?
4. Conscientiousness (CON) - Is the person efficient and organized versus sloppy?
5. Openness (OPN) - Is the person inventive and curious versus dogmatic?

This project is a focused application of automatic personality detection and the results of this project may be used to segment movie characters in various categories, and then perform an in depth character analysis of a movie or a series of movies. More generally, this may further be used to gain some insight into human personalities based on the way they converse.

3) Background

We first came up with the idea of analyzing a movie character based on the entire movie or series. In the process of looking for resources to address our problem statement, we realized the different issues we would encounter and so we made our problem statement simpler and more precise by brainstorming and researching. In this process, we found the following links helpful.

A. Film Script Analysis. https://github.com/AdeboyeML/Film_Script_Analysis - This was one of our first inspirations for our project idea. The author did an analysis on characters, scenes, character's interaction with other characters, etc to provide insights into different movies

B. Deep Learning-Based Document Modeling for Personality Detection from Text. deep-learning-based-personality-detection.pdf (sentic.net) : This paper presents a novel approach to automatic personality recognition using pre-trained contextual embeddings (BERT and RoBERTa) and attentive neural networks. Their models improve the state-of-art results on the monologue Essays dataset by 2.49%, and establish a solid benchmark on FriendsPersona dataset. The results of this model are a **benchmark** for our implementation

C. Personality Prediction from Text. gorkemgoknar/personality-detection-text - githubmemory : This paper compares models like Random Forest, SVM, Decision Tree, Naive Bayes, Logistic Regression on essays dataset.

We are building up on the code base mentioned in parts A and D of this section. We used part A for preprocessing of the movie scripts data and part D for training the models.

4) Summary of Our Contributions

1. Contribution(s) in Code: We want to try different word embedding techniques like the ones obtained from Glove and BERT while getting a feature representation for our input text.
2. Contribution(s) in Application: OCEAN trait prediction is more common in the image domain and mainly on social media datasets in the text-domain. In this project, we want to perform it on an entirely different application. We are interested in gauging the personality trait of a character in a particular scene based on his dialogues.
3. Contribution(s) in Data: N/A
4. Contribution(s) in Algorithm: N/A
5. Contribution(s) in Analysis: N/A

5) Detailed Description of Each Proposed Contribution

5.1 Methods

Predicting OCEAN labels in an unsupervised manner is not very common. Therefore, we decided to use annotated moviescript datasets. We found one which is based on FRIENDS TV series scripts [Jinho D. Choi, 2020], but the dataset size is relatively very small (~750 samples) and we feel that it might not be sufficient. So, instead, the backbone idea here is to first get some pre-trained parameters for our supervised classification problem by training it on other datasets [Fabio Celli, 2013] [Jason, 2018] (the two annotated datasets based on essays and social media platforms mentioned in the datasets sections earlier), which gives some context to the model before finally training it on the scripts dataset. So, the latter step here helps the model to tune itself to the movie scripts context which is the transfer learning part.

5.2 Experiments

Model Experiment 1:
We want to evaluate the performance of models like SVM, Decision Tree, Naive Bayes, Random forest, Logistic regression, and XGBoost for our OCEAN label classification problem [Gorkem, 2020]. We will extract features from text to vectorize the data with GloVe approach. We want to compare these models with a simple multi-layer perceptron (MLP) which has 3-4 layers. Each model refers to five binary classifiers for each of the OCEAN labels. We will tune the hyperparameters of each model and we will select the best model based on the performance and predicted accuracies.
Model Experiment 2:
Word embeddings are used to represent the text as a densely distributed vector. A good word embedding model makes sure that similar texts or texts which are based out of similar context have their vector representations close to each other in the high dimensional space. Therefore, high-quality word embeddings are essential for our model to get good context about the input text. So, we want to test another approach like Bert, which is known to extract contextual information from a text, to extract word embeddings and finalize the one which gives the best performance.

Datasets:
1. Essays - http://web.archive.org/web/20160519045708/http://mypersonality.org/wiki/doku.php?id=wcpr13: This has 2469 essays which have around 600 words each as input text (X).
2. Mypersonality dataset - https://github.com/jcl132/personality-prediction-from-text/tree/master/data/myPersonality: This has 9918 facebook statuses as input text (X).

3. FRIENDS TV series dataset - https://github.com/emorynlp/personality-detection: This has 712 concatenated set of dialogues of character in a scene as input text (X).

For all the above three datasets, the input is text the corresponding labels to be predicted (y) are in the form where one or more of the five categories has a binary yes (1) and the remaining categories are labeled no (0).

Performance Metrics and Train-Test Split:
As it is a classification task, we plan to use Accuracy on test data (20% of the whole data) as our performance metric.

Dataset Experiment 1:
Train the model directly with the FRIENDS dataset [Jinho D. Choi, 2020] and measure the accuracy of the test data which is plain supervised learning.

Dataset Experiment 2:
We'll perform transfer learning here. First, train the model of the other two datasets based on essays and social media [Fabio Celli, 2013] [Jason, 2018] and then on the FRIENDS dataset and measure the accuracies. Measure the impact of pretraining the model on each of the other two datasets before training on our final scripts data. For the models which use sklearn, it is not possible to finetune the parameters. In such a case, we combined the datasets using pandas and trained the model on this combined dataset. This can be done in the following way:
Measure the test accuracies after training on
a) Essay dataset b) Social Media dataset c) Friends dataset d) Essay and Social Media dataset e) Essay, Social Media and Friends datasets
Differences between the above accuracies will help us evaluate the importance of each dataset to our final model.

Final Testing Model:
We Preprocessed the dataset of the movie 'The Social Network' to segment it based on scenes, concatenated all the dialogues of a character in a particular scene which is going to be the text input to the final model we choose based on the accuracy scores.

5.3) Results

Dataset 1 = Essays; Dataset 2 = Social Media; Dataset 3 = F.R.I.E.N.D.S dialogues

	Dataset 1					Dataset 2					Dataset 3					Dataset 1 + 2					Dataset 1 + 2 + 3				
Glove embedding	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**
SVM	**0.56**	**0.61**	**0.54**	**0.56**	0.6	**0.57**	**0.64**	0.54	0.54	0.76	0.64	0.52	0.56	0.53	0.62	0.54	0.6	0.54	0.54	0.7	0.51	0.57	0.51	0.51	0.68
Decision Tree	0.53	0.52	0.51	0.51	0.52	0.52	0.54	0.52	0.51	0.64	0.52	0.47	0.5	0.44	0.53	0.53	0.54	0.51	0.52	0.6	0.48	0.49	0.46	0.46	0.56
Naive Bayes	0.48	0.57	0.54	0.53	0.57	0.49	0.59	0.54	**0.56**	0.6	0.6	0.53	0.47	**0.58**	0.56	0.51	0.52	0.54	0.49	0.49	0.3	0.32	0.29	0.28	0.32
Logistic Regression	0.55	0.6	0.53	0.55	0.6	0.57	0.63	0.54	0.55	0.76	0.6	0.56	0.58	0.52	0.58	0.54	0.59	0.54	0.55	0.69	0.51	0.57	0.51	0.52	0.68
Random Forest	0.48	0.53	0.54	0.53	0.57	0.56	0.62	0.55	0.55	**0.76**	0.59	0.53	0.52	0.43	0.57	0.56	0.6	0.54	0.55	0.7	0.53	**0.58**	0.53	0.52	0.68
XGBoost	0.53	0.53	0.54	0.55	0.55	0.55	0.6	0.54	0.54	0.74	0.56	**0.57**	**0.56**	0.48	0.52	0.55	**0.59**	**0.54**	0.53	0.69	0.53	0.56	0.53	0.51	**0.68**
MLP	0.54	0.58	0.53	0.52	**0.61**	0.56	0.61	**0.56**	0.55	0.75	**0.66**	0.55	0.54	0.45	**0.57**	**0.56**	0.58	0.53	**0.56**	**0.69**	**0.56**	0.55	**0.54**	**0.55**	0.57

Figure 1: Accuracy scores of various ML models using Glove embeddings.

	Dataset 1					Dataset 2					Dataset 3					Dataset 1 + 2					Dataset 1 + 2 + 3				
BERT embedding	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**
SVM	**0.55**	0.56	0.56	0.58	0.56	0.58	0.62	0.55	0.59	0.76	0.58	**0.5**	0.59	0.55	0.58	0.58	0.59	0.56	0.57	0.7	0.58	0.59	0.55	0.56	0.71
Decision Tree	0.52	0.51	0.5	0.48	0.53	0.55	0.55	0.54	0.52	0.63	0.58	0.5	0.57	0.46	0.54	0.53	0.53	0.51	0.53	0.58	0.5	0.52	0.48	0.48	0.56
Naive Bayes	0.52	0.55	0.54	0.57	0.56	0.5	0.6	0.52	0.54	0.49	**0.64**	0.47	0.53	0.51	0.58	0.51	0.58	0.52	0.55	0.59	0.38	0.46	0.42	0.41	0.41
Logistic Regression	0.54	**0.57**	0.54	0.55	0.57	0.58	0.62	0.56	0.6	0.75	0.61	0.45	0.6	0.55	0.6	0.57	0.6	0.56	0.59	0.69	0.58	0.6	0.56	0.57	0.7
Random Forest	0.52	0.53	0.53	0.53	0.58	0.59	0.62	0.56	0.58	0.76	0.6	0.46	0.56	0.5	0.64	0.56	0.6	0.58	0.58	**0.71**	0.56	0.58	0.55	0.54	0.69
XGBoost	0.51	0.55	0.56	0.55	0.55	**0.61**	**0.64**	**0.57**	**0.59**	**0.77**	0.6	0.4	0.53	0.53	0.61	0.57	**0.61**	0.57	**0.6**	0.7	0.57	**0.59**	0.56	**0.57**	0.7
MLP	0.54	0.54	**0.57**	**0.58**	**0.58**	0.59	0.62	0.56	0.59	0.73	0.6	0.45	**0.6**	**0.56**	**0.64**	**0.59**	0.6	**0.58**	0.59	0.7	**0.62**	0.57	**0.62**	0.55	**0.71**

Figure 2: Accuracy scores of various ML models using BERT embeddings.

		Final Movies Dataset - THE SOCIAL NETWORK					
Character	**Scene**	**Dialogues**	**EXT**	**NEU**	**AGR**	**CON**	**OPN**
MARK	1	How do you distinguish yourself in a population of people who all got 1600 on their S	0	0	1	0	1
ERICA	1	I didn't know they take SAT's in China. You got 1600? I like guys who row crew. Have y	0	0	1	0	1
BRICA	1	Does that mean that you actually got nothing wrong?	0	0	0	0	1
MAFT	1	Wait, wait, thisi reel?	0	0	0	0	1
PAUL YOUNG	11	T.T'S A GOOD THING YOU DON'T HAVE A BUSFARE IT WOULD FALL THROUGH A HOLE IN YOUR POCKET AND YOU'D LOSE IT IN THE SNOW ON THE WHERE YOU CAN LIVE IN THE LOVE OF THE	0	1	0	0	1
MARK	11	Adams has no security but limits the Leverett is a little better. It's slightly obnoxious t	0	0	1	0	1
EDUARDO	11	What's going on? Mark. What's going on? 15, Did you and Erica split up? It's on your blog. Are you alright? I am here for you. Are you okay? Why? You think that's such a good idea?	0	0	0	0	1
MARY	11	What's the algozithm?	1	0	1	0	1
EDUARDO	88	It's time to monetize the thing. Did you hear what I said? I said it's time to monetize t	0	0	1	1	1
MARK	88	Those Asian girls were cute. What were their names? When? What does that mean?	0	0	1	0	1
MARX	88	The girls. EDUARDO's speed reading the letter. I know what it says.	0	1	1	0	1
GAGE	88	When we met in January, I expressed my doubts about the sitewhere it stood with graphics, how much programming was left that I had not anticipated	0	0	0	0	1
JENNY	103	They are not gonna card us here. Look around. Tell him they are not gonna card us. I think Wardo's jealous.	0	1	0	0	1
EDUARDO	103	They might. It' l.l be embarrassing, MARK Unless you are the Ballet Theatre of Hartfor	0	1	1	0	1
MARK	103	They are not gonna card us. Are you gonna talk about ads again? This isn't a business. He's a god, he can be as late as he wants. What is he?	0	0	1	0	1
EDUARDO	104	I honestly wasn't jealous. I was nervous. I didn't know him at all but I had done an internet search and asked around. He struck me as kind of a wild man. TNT. 66 NIGHT He crashed out of two pretty big internet companies inspectacular fashion and he's got a reputation with drugs. We don't need him. And he does own a watch. SEAN stops at a table to shake hands with a guy in a suit and kiss his girlfriend. It's sort of an incongruous sightthis 22 year old kid who's able to work a room like Sinatra. Who the hell is this? Take your time... We were waiting for	0	0	0	0	1
GRETCHEN	104	Why?	0	0	0	0	1
JENNY	104	Why? Stop it. SEAN makes his way'over to MARK's table An appletini?	0	0	1	0	1
MARK	104	He also founded'the companies. He's here. SEAN PARKER has stepped into the restaurant and is. saying hello to the hostess while hugging a waitress. Great to meet you.	1	0	1	1	1
SEAN	104	I am Sean Parker. EDU.ARDO How do you do. You must be Eduardo, And Jenny. And	0	1	1	0	1

Figure 3: OCEAN label predictions on our scene-segmented movies dataset for each character

From the above results in Figure 2 and 3, we can mainly understand the impact of the amount of data, quality of data and the embedding model we use on the performance of a particular ML model. The results are on-par or better than the benchmark results which the github author used for essays dataset. Since Bert embeddings consider the context of where a particular word occurs while generating an embedding, it’s supposed to perform better for our task which matches with what we have observed. Since Dataset1 has only 2469 essays, out of which 80% are used for training, though the data is rich in itself, it is limited by the number of datapoints which results in a poor performance of all our ML models. Dataset2 is not as rich as dataset1, but it's relatively huge (9918 samples) and therefore the accuracy has improved. XGBoost and MLP models tend to perform better with more data. Dataset 3's context is the closest to what we want to achieve in the end, but it's a very small database (712 samples). But, the models did perform reasonably well. Finally when we tried combinations of datasets, we expected the accuracies to go up, but they either increased a little or stayed the same. We also observed that models like Naive bayes, Decision tree performed worse when all the three datasets (and therefore contexts) were mixed up together. In this analysis, we used a simple MLP model to match the complexity with the remaining traditional ML models. If we had used a more complex RNN, we expect the transfer learning to give better results as it has more parameters to train and therefore better scope of understanding the context of a text. Figure 3 shows the final predictions on un-labelled scene-segmented movie data.

6) Compute / Other Resources Used

We used google Colab for preprocessing the datasets,and training all our models, predicting the results.

7) Conclusion

We successfully predicted the Big 5 personality traits of movie characters for each scene using the film script. We learned how to apply machine learning models to a different domain and how this affects data preprocessing, embedding generation and choice of model. Scene wise personality detection of characters from movie scripts can be used in various domains. One of them can be : Movie directors can use this to see which personality suits an actor/actress the best by analyzing the ratings of that movie and personality of that actor/actress as predicted by our model.

Due to the domain shift we saw that none of our models were working very well in the beginning. After a few brainstorming sessions among ourselves and with our project mentor(TA) we realised that in this domain, the context is quite important and hence moved to BERT embedding from Glove. Even after this the results didn't improve significantly and we realised the size of dataset and lack of domain specific labelled dataset was causing the hindrance of improved results. If we had a huge dataset, whose context aligns with our end goal, we could've achieved a much better model.

This is slightly a sensitive topic raising multiple ethical questions. These ethical implications are listed in a paper [10] which states that: Algorithmic bias could be an issue in this context, making certain groups susceptible to negative effects of psychological targeting. Various research has suggested that detecting private characteristics of users can lead to biases, and put certain groups in unfair situations.

Other Prior Work / References:

1. Automatic Text-Based Personality Recognition on Monologues and Multiparty Dialogues Using Attentive Networks and Contextual Embeddings. https://ojs.aaai.org/index.php/AAAI/article/view/7182/7036 : In this article, the 4 CIS 519 Project Proposal author presents a method to extract personality traits from stream of-consciousness essays using a convolutional neural network (CNN)
2. Using textual data for Personality Prediction:A Machine Learning Approach. https://www.researchgate.net/publication/339977005_Using_textual_data_for_Personality_PredictionA_Machine_Learning_Approach : This uses classification algorithms like AdaBoost, Multinomial Naïve Bayes and LDA to their twitter dataset, to classify personality in any one of the five class labels provided by Big five test of psychology.
3. TwitPersonality: Computing Personality Traits from Tweets Using Word Embeddings and Supervised Learning. https://www.mdpi.com/2078-2489/9/5/127/pdf : This approach is based on transfer learning, namely modeling an algorithm for a context and applying it on a (slightly) different, yet related, context. They derived a machine learning model for predicting personality of Facebook users based on their status updates, then tested it on Twitter users' tweets.
4. A Survey of Automatic Personality Detection from Text. https://aclanthology.org/2020.coling-main.553.pdf : Summarizes all the prior work done in the field of personality detection using text. This helped us get the gist of the progress so far in this domain.
5. PersonalityInsightsV3 - http://watson-developer-cloud.github.io/node-sdk/master/classes/personalityinsightsv3.html . For personality profiling: IBM Watson has a library called PersonalityInsightsV3. We hope to produce something similar to this, and also give OCEAN model of personality (i.e. 5 personality traits)
6. https://www.researchgate.net/publication/342656426_A_Systematic_Literature_Review_of_Personality_Trait_Classification_from_Textual_Content
7. Recent trends in deep learning based personality detection | SpringerLink
8. Personality Classification from Online Text using Machine Learning Approach (thesai.org)
9. https://www.storyfit.com/blog/script-characters-ai.
10. A Survey of Automatic Personality Detection from Text. https://aclanthology.org/2020.coling-main.553.pdf

Broader Dissemination Information:

Your report title and the list of team members will be published on the class website. Would you also like your pdf report to be published?
YES

If your answer to the above question is yes, are there any other links to github / youtube / blog post / project website that you would like to publish alongside the report? If so, list them here.

PERSON (S)	TASK (S)	Wk5				Wk6				Wk7				Wk8				Wk9			
		OCT																	NOV		
		S 3	M 4	W 6	Th 7	S 10	M 11	W 13	Th 14	S 17	M 18	W 20	Th 21	S 24	M 25	W 27	Th 28	S 31	M 1	W 3	Th 4
Shivani, Jay, Saurabh	More Literature Review and finalizing a model																				
Shivani, Jay	Implement the emotion analysis model																				
Jay, Saurabh	Analyze the results, improve them					●															
Saurabh, Shivani, Jay	Analyze each part of the code, improve aspects like loss function, optimizers etc.																				

PERSON (S)	TASK (S)	Wk10				Wk11				Wk12				Wk13				Wk14			
		NOV																	DEC		
		S 3	M 4	W 6	Th 7	S 10	M 11	W 13	Th 14	S 17	M 18	W 20	Th 21	S 24	M 25	W 27	Th 28	S 31	M 1	W 3	Th 4
Shivani, Jay, Saurabh	Implement the emotion analysis model																				
Shivani, Jay, Saurabh	Analyze the results, improve them																				
Shivani, Jay, Saurabh	Analyze each part of the code, improve aspects like loss function, optimizers, embeddings etc.																				

SRS_CIS-519-401 2021C Project Report 2 (Midway Progress Report)

Jay Gala, Saurabh Raut, Shivani Rapole

TOTAL POINTS

7 / 7

QUESTION 1

1 Does the provided template including the 4-page limit (excluding exempted portions), with reasonable responses to all questions? **1 / 1**

✓ **- 0 pts Correct**

QUESTION 2

2 Has feedback from the last round been effectively addressed? **2 / 2**

✓ **- 0 pts Correct**

QUESTION 3

3 Has the team identified a clear topic and viable new target contribution, as per the project specifications provided in class? **1 / 1**

✓ **- 0 pts Correct**

QUESTION 4

4 Has the team moved in a non-trivial way towards their target contribution? **1 / 1**

✓ **- 0 pts Correct**

1 You didn't include the progress towards the final goal in this section, but since you did restructure your projet it seems like it is still feasible.

QUESTION 5

5 Has a clear and systematic work plan been formulated for the remaining weeks? **2 / 2**

✓ **- 0 pts Correct**

CIS 519 Project Proposal

Emotion Analysis of Movie Characters based on Movie Scripts

Team: Saurabh Raut, Jay Gala, Shivani Reddy Rapole.
Project Mentor TA: Pratik Kunapuli

1) Introduction

Problem statement:
Our main goal is to understand the emotion of movie characters based on movie scripts. We plan to perform a scene wise personality trait prediction depending on the dialogues of the character in that particular scene. So, the input to our supervised multi-class classification model is the set of dialogues of a character in a scene and the output is one of the Big Five traits (OCEAN model) - openness, conscientiousness, extraversion, agreeableness, and neuroticism.

Training Data:
One of the main issues we faced in this process was finding a movie or TV series scripts dataset which is annotated with OCEAN model traits. We found one but it was a small dataset, which needs to be further partitioned into training and testing data. So, instead of making our model entirely dependent on this data, we want to first make it learn classifying into the traits based on two other annotated datasets from platforms like Twitter and general essays. We then want to adapt these pretrained parameters to the dialogue scripts domain by training and therefore tuning it to our FRIENDS TV series script dataset.

Dataset Links:

1. http://web.archive.org/web/20160519045708/http://mypersonality.org/wiki/doku.php?id=wcpr13
2. https://github.com/jcl132/personality-prediction-from-text/tree/master/data/myPersonality
3. https://github.com/emorynlp/personality-detection

Testing Data:
We have two options for testing our final model. The primary one is obtained by splitting the earlier FRIENDS dataset which is annotated. So, accuracy and F1 score can be used as metrics to evaluate the performance of our model. The secondary dataset is optional. It contains various movie scripts which are not annotated, we will manually validate the predicted results for a few samples and comment on the performance of our model.

CIS 519 Project Proposal

Dataset links:

1. https://github.com/emorynlp/personality-detection
2. https://github.com/AdeboyeML/Film_Script_Analysis

Motivation:

This project is aimed at giving an insight into a movie's characters' personality and how their personality or emotion (short term personality) is changing over the course of the movie. This will be done just based on the dialogues a particular character says and the scenes in which they are involved. We will analyze the character's personality using the famous Big 5 Personality Trait model (OCEAN model) for each scene they are in, and then, evaluate the change of emotions or personality in each scene throughout the movie.

This is a focused application of automatic personality detection and the results of this project may be used to segment movie characters in various categories, and then perform an in depth character analysis of a movie or a series of movies. More generally, this may further be used to gain some insight on human's personalities based on the way they converse. However, this is slightly a sensitive topic raising multiple ethical questions.

These ethical implications are listed in the following paper - (A Survey of Automatic Personality Detection from Text. https://aclanthology.org/2020.coling-main.553.pdf). This paper states that: "Algorithmic bias could be an issue in this context, making certain groups susceptible to negative effects of psychological targeting. Various research has suggested that detecting private characteristics of users can lead to biases (Bolukbasi et al., 2016), and put certain groups in unfair situations. The use of personality detection via psychographic targeting in advertising and marketing also brings ethical concerns. However, psychographic targeting can be particularly harmful for vulnerable groups who engage with risky behaviors, such as targeting addicted people with online gambling advertising (Matz et al., 2017; Gladstone et al., 2019)."

After going through these ethical implications, we decided to restrict ourselves to movie characters which are not real humans.

2) How We Have Addressed Feedback From the Proposal Evaluations

Feedback on Proposal 1:

CIS 519 Project Proposal

1 All evaluation **3 / 3**

✓ + 1 pts **Team formation**

✓ + 1 pts **Understanding of source projects, identification of clear target contribution**

✓ + 1 pts **Followed template, with reasonable responses to all or most questions**

+ 0 pts Does this proposal include a code contribution?

✓ + 0 pts **Does this proposal include an application contribution?**

✓ + 0 pts **Does this proposal include a data contribution?**

+ 0 pts Does this proposal include an algorithm contribution?

✓ + 0 pts **Does this proposal include a technique / analysis contribution?**

+ 0 pts (TAs, don't remove this item.. this is a set of things to consider when adding free-form comments)

- Is this application of ML well-motivated i.e., does this problem need ML?
- Have they left out related work you are aware of?
- Are there ethical concerns?
- Is this project feasible given the time & compute limitations, and team expertise?
- Is the risk mitigation plan good enough, does it aim to build up a minimum viable product quickly?

Overall, good proposal. My biggest concern with this project is the lack of foundational work done ahead of you since you're applying some existing work to an entirely new domain. Additionally, it seems like the dataset you intend to work with requires a lot of pre-processing and manual annotation. Finally, the expected challenges you mention are fairly significant but you don't really mention any way of handling them. My general thoughts on this is if you're going to be using your friends to annotate this data, it's a lot easier to have them come up with OCEAN labels for the characters in each scene. Then you can learn how to predict those personality traits from the dialogue in the scene. Otherwise, it seems like an interesting project and I look forward to seeing the progress.

-Pratik

1. Annotating the labels here seems like a non-trivial amount of work.
2. How do you intend to pre-process this data into something you can run an unsupervised algorithm on?
3. Are these big 5 personality traits going to be the output of the model? How are these labeled and measured?
4. How do you intend to address this specifically?

Addressing the above issues:
1 - Yes, and we realised finding an unsupervised model for our task is unlikely. So, we changed the dataset and found an annotated one.
2 - This is addressed in the previous point.
3 - Yes, in the current supervised model, Big 5 traits are the output. The dataset is annotated that way too.
4 - Instead of either annotating and pre-processing the data or using an unsupervised approach, we switched to supervised learning with an annotated dataset.

CIS 519 Project Proposal

We also interacted with our mentor Pratik about this, his guidance has helped us lay out our plan wisely. He suggested that a supervised learning approach makes more sense if we are using text as data and then producing labels should likely be supervised from some labeled data which we also agreed with after some research. We incorporated his suggestions and switched to a supervised learning approach and searched for an annotated dataset.

3) Prior Work We are Closely Building From

We first came up with the idea of analysing a movie character based on the entire movie or series. In the process of looking for resources to address our problem statement, we realized the different issues we would encounter and so we made our problem statement simpler and more precise by brainstorming and researching. In this process, we found the following links helpful.

A. Film Script Analysis. https://github.com/AdeboyeML/Film_Script_Analysis - This was one of our first inspirations for our project idea. The author did an analysis on characters, scenes, character's interaction with other characters, etc to provide insights into different movies
B. How to use StoryFit AI to analyze movie characters. https://www.storyfit.com/blog/script-characters-ai. This is more in line with what we want to achieve. The author did a major character breakdown based on various factors and created an emotional palette for each character to generate a final personality profile.
C. PersonalityInsightsV3. http://watson-developer-cloud.github.io/node-sdk/master/classes/personalityinsightsv3.html. For personality profiling: IBM Watson has a library called PersonalityInsightsV3. We hope to produce something similar to this, and also give OCEAN model of personality (i.e. 5 personality traits)
D. Deep Learning-Based Document Modeling for Personality Detection from Text. deep-learning-based-personality-detection.pdf (sentic.net) : This paper presents a novel approach to automatic personality recognition using pre-trained contextual embeddings (BERT and RoBERTa) and attentive neural networks. Their models largely improve the state-of-art results on the monologue Essays dataset by 2.49%, and establish a solid benchmark on FriendsPersona dataset. The results of this model are a **benchmark** for our implementation
E. Automatic Text-Based Personality Recognition on Monologues and Multiparty Dialogues Using Attentive Networks and Contextual Embeddings. https://ojs.aaai.org/index.php/AAAI/article/view/7182/7036 : In this article, the

CIS 519 Project Proposal

author presents a method to extract personality traits from stream of-consciousness essays using a convolutional neural network (CNN)

F. Personality Prediction from Text. gorkemgoknar/personality-detection-text - githubmemory : This paper compares models like Random Forest, SVM, Decision Tree, Naive Bayes, Logistic Regression on essays dataset.

G. Using textual data for Personality Prediction:A Machine Learning Approach. (PDF) Using textual data for Personality Prediction:A Machine Learning Approach (researchgate.net) : This uses classification algorithms like AdaBoost, Multinomial Naïve Bayes and LDA to their twitter dataset, to classify personality in any one of the five class labels provided by Big five test of psychology.

H. TwitPersonality: Computing Personality Traits from Tweets Using Word Embeddings and Supervised Learning. https://www.mdpi.com/2078-2489/9/5/127/pdf : This approach is based on transfer learning, namely modeling an algorithm for a context and applying it on a (slightly) different, yet related, context. They derived a machine learning model for predicting personality of Facebook users based on their status updates, then tested it on Twitter users' tweets.

I. A Survey of Automatic Personality Detection from Text. https://aclanthology.org/2020.coling-main.553.pdf : Summarizes all the prior work done in the field of personality detection using text. This helped us get the gist of the progress so far in this domain.

4) What We are Contributing

1. Contribution(s) in Code: We want to try different word embedding techniques like the ones obtained from BERT and RoBERTa while getting a feature representation for our input text.
2. Contribution(s) in Application: OCEAN trait prediction is more common in the image domain and mainly on social media datasets in the text domain. In this project, we want to perform it on an entirely different application. We are interested in gauging the emotional state of a character in a particular scene based on his dialogues.
3. Contribution(s) in Data: N/A
4. Contribution(s) in Algorithm: N/A
5. Contribution(s) in Analysis: N/A

CIS 519 Project Proposal

5) Detailed Description of Each Proposed Contribution, Progress Towards It, and Any Difficulties Encountered So Far

5.1 Methods

Our main goal is to predict the big five traits (OCEAN labels) of a particular character in a scene which are:

1. Extroversion (EXT) - Is the person outgoing, talkative, and energetic versus reserved and solitary?
2. Neuroticism (NEU) - Is the person sensitive and nervous versus secure and confident?
3. Agreeableness (AGR) - Is the person trustworthy, straightforward, generous, and modest versus unreliable, complicated, meager, and boastful?
4. Conscientiousness (CON) - Is the person efficient and organized versus sloppy and careless?
5. Openness (OPN) - Is the person inventive and curious versus dogmatic and cautious?

Predicting OCEAN labels in an unsupervised manner is not very common. Therefore, we decided to use annotated moviescript datasets. We found one which is based on FRIENDS TV series scripts [Jinho D. Choi, 2020], but the dataset size is relatively very small (~750 samples) and we feel that it might not be sufficient. So, instead, the backbone idea here is to first get some pre-trained parameters for our supervised multi-class classification problem by training it on other datasets [Fabio Celli, 2013] [Jason, 2018] (the two annotated datasets based on essays and social media platforms mentioned in the datasets sections earlier), which gives some context to the model before finally training it on the scripts dataset. So, the latter step here helps the model to tune itself to the movie scripts context which is the transfer learning part.

5.2 Experiments and Results

Model Experiment 1:

We want to evaluate the performance of models like Random forest, logistic regression for our multi-class classification problem [Gorkem, 2020]. We will extract features from text to vectorize the data with bags of words and GloVe approach. We want to compare these models with the second model which is based on a 7-layer Convolutional Neural Network (CNN) [Erik cambria, 2017]. We will select the best model based on the performance and predicted accuracies. For the above models, we will use k-fold cross validation to evaluate the performance of the model while training.

Model Experiment 2:

CIS 519 Project Proposal

Word embeddings are used to represent the text as a dense distributed vector. A good word embedding model makes sure that similar texts or texts which are based out of similar context have their vector representations close to each other in the high dimensional space. Therefore, high quality word embeddings are essential for our model to get good context about the input text. So, we want to test various approaches like word2Vec, Bert, etc to extract word embeddings and finalize the one which gives the best performance.

Datasets:

1. Essays - http://web.archive.org/web/20160519045708/http://mypersonality.org/wiki/doku.php?id=wcpr13: This has essays which have around 600 words each as input text (X) and the labels to be predicted (y) are in the form where one or more of the five categories has a binary yes (1) and the remaining categories are labelled no (0).
2. Mypersonality dataset - https://github.com/jcl132/personality-prediction-from-text/tree/master/data/myPersonality: This has facebook statuses as input text (X) and the labels to be predicted (y) are in the form where one or more of the five categories has a binary yes (1) and the remaining categories are labelled no (0).
3. FRIENDS TV series dataset - https://github.com/emorynlp/personality-detection: This has a concatenated set of dialogues of a character in a particular scene as input text (X) and the corresponding labels to be predicted (y) are in the form where one or more of the five categories has a binary yes (1) and the remaining categories are labelled no (0).

Performance Metrics:

As it is a classification task, we plan to use Accuracy and F1 score as our performance metric.

Dataset Experiment 1:

Train the model directly with the FRIENDS dataset [Jinho D. Choi, 2020] and measure the accuracy of the test data which is plain supervised learning.

Dataset Experiment 2:

We'll perform transfer learning here. First train the model of the other two datasets based on essays and social media [Fabio Celli, 2013] [Jason, 2018] and then on the FRIENDS dataset and measure the accuracies. Measure the impact of pretraining the model on each of the other two datasets before training on our final scripts data. This can be done is the following way:

Measure the test accuracies after training on

a) Essay dataset

CIS 519 Project Proposal

b) Social Media dataset
c) Friends dataset
d) Essay and Friends dataset
e) Social Media and Friends dataset
f) Essay, Social Media and Friends datasets

Differences between the above accuracies will help us evaluate the importance of each dataset to our final model.

6) Risk Mitigation Plan

- So first we are going to build a minimum viable project as follows: Generate word embeddings using BERT and RoBerta for the three datasets mentioned above. Train a CNN and Random forest Classifier using these embeddings and predict the OCEAN personality traits on movie scripts.
- Yes, first we are planning to start with a simplified setting of using the word2Vec embeddings and training it simply on CNN and then use other embeddings and other models.
- So if our approach doesn't work, we would first try to find the reason behind failure. So even if we are not able to solve the reason causing failure, we can make a useful report by establishing that using this model or approach does not generate good results on this data because of the failure reason we found out.
- If we find there is too much to compute then we can train our models on possibly different subsets of data so as to reduce input size. Also we can reduce the size of each input by different approaches of dimensionality reduction like PCA.
- We had decided that we would try our algorithm on some simple dialogues like "Yes, I agree with you all" and check the probabilities generated for specific classes like "agreeableness" for the example above.
- We will definitely try to evaluate the reason for our failure. Looking at the number of datasets we are going to train on and the number of models we are planning to train on these datasets it is assured that some of them will fail. We will analyze the dataset and model after each train-test cycle by analysis of that data and different performance metrics for that model like precision etc. This will help us improve on the data as well as the model in each step.