WGU Capstone Project

TWITTER PROFILE ANALYZER

JOSHUA DIX

Table of Contents

Letter of Transmittal

Joshua A. Dix
1165 E 1080 S
#214
Provo, UT 84606

Ashley Hollings
CEO
3Delta Consulting Incorporated
2675 W 600 N
Suite 200
Lindon, UT 84606

Dear Ms. Ashley Hollings,

Due to the near worldwide popularity of social media, Twitter has become a modern-day public square filled with voices and viewers from all over the globe. This far-reaching and yet less personal form of public discourse has led some to participate in cyberbullying; specifically, discrimination and hate speech. Unfortunately, it was brought to our attention that one of our employees was tweeting such cognate content. This situation brought the ethics of our company into question and concern that clientele may have been subject to similar biases. To rectify this situation the said employee was terminated, and a public apology was issued.

To preserve the integrity of 3Delta Consulting and to avoid a future repeat of this situation further steps are needed. I believe the solution is to be proactive in our future hires and retainment of current employees through a machine learning social media analyzer that is capably unbiased in analyzing applicant and current employee's tweets for hate speech before harming public relations and our client relationships.

The proposed social media analyzer system supports our cultural diversity and inclusion efforts. Since the insights from the data visualizations would be available through a dashboard, it would aid in the future employee hiring process as well as the continual prevention of discrimination within our company.

The funding required to implement this analyzer is approximately $50,000. Most of the funding would be allocated for development and human resource-related costs. It also includes any in-scope development time with the project. This project should take an estimated twelve weeks to complete. I am confident that my knowledge of python and previous development experience, along with my cross-functional team experience will allow me to successfully lead a team and bring this solution to fruition.

If you have any questions regarding the attached report, please contact me at (801) 919-6457.

Sincerely,

Joshua A. Dix
VP of Research and Development

Project Recommendation

Problem Summary

3Delta Consulting Inc. had an employee that engaged in hate speech on Twitter. Unfortunately, this resulted in the termination of said employee, and damaged relationships with clientele over concerns of bias. To remedy this situation 3Delta Consulting Inc. issued a public apology but would like to proactively prevent a repeat. To do this 3Delta Consulting Inc. would like to use a machine learning tweet analyzer to help vet future applicants and to periodically monitor current employees' tweets as part of their continued employment.

Application Benefits

The machine learning tweet analyzer will allow 3Delta Consulting to root out future applicants that participate in hate speech thus helping them hire employees who are more compatible with the desired company culture and mission. It will also help ensure that current employees do not cause harm to the company-client relationship by tweeting hate speech. The machine learning analyzer is also unbiased, much quicker, and more thorough in its analysis than a human would be.

Application Description

The machine learning tweet analyzer will be developed using Python and their data analysis libraries. The program will use Google Colab, which is an online interactive data science environment like a Jupyter Notebook that allows collaboration over the web. Google Colab has the capabilities of allowing you to separate code and text into individual cells, this allows you to notate what you are trying to accomplish, along with incremental development. The program will then use a Python Library converter tool to create a JSON file that describes the model including its setup, topology, type of layers, inputs, outputs, etc., and a binary file that contains the trained weights. This will then be connected to a custom UI created with React.js (an Opensource frontend JavaScript framework), and D3.js (an Opensource framework for data analytics).

The machine learning tweet analyzer will aid 3Delta Consulting Inc. in hiring and monitoring employees that meet its desired company culture and mission. It will also aid in keeping positive business-client relationships by including data visualizations of the tweets from the applicant's/employee's accounts.

Data Description

The dataset used to create the machine learning tweet analyzer prototype was acquired from Kaggle.com. The dataset contains information on 24,783 tweets. The data contains the following:

- Count – number of CrowdFlower users who coded each tweet (min is 3, sometimes more users coded a tweet when judgments were determined to be unreliable by CF)
- Hate Speech – number of CF users who judged the tweet to be hate speech
- Offensive Language - number of CF users who judged the tweet to be offensive
- Neither – number of CF users who judged the tweet to be neither offensive nor hate speech
- Class – label for majority of CF users. (0 – hate speech, 1 – offensive language, 2 – neither)
- Tweet – The tweet from the text

The data uses a CSV format that will be used to create the machine learning tweet analyzer. Before the data can be used, it will need to be preprocessed and normalized so the machine learning algorithm can use it. The data will also need to be analyzed to eliminate unnecessary attributes; this should increase accuracy and response time. The main limitation of this data set is that it is skewed toward tweets that are

classified as offensive language. The machine learning tweet analyzer should be able to adapt to any number of tweets.

Objective & Hypotheses

The objective of the machine learning tweet analyzer is to help 3Delta Consulting Inc. hire and monitor employees' tweets so that 3Delta Consulting may prevent situations that would damage the company's relations with its clients and must terminate current employees. The machine learning tweet analyzer hypothesizes that 3Delta Consulting Inc. can mitigate some business-client damage while increasing its brand recognition through having a great company culture.

Methodology

The development methodology that will be used to develop the machine learning tweet analyzer is the agile methodology. This methodology was chosen because it allows for a high level of stakeholder engagement with the development team, along with early and predictable deliveries. The agile methodology is based on iterative cycles called sprints. This allows for increased flexibility and quality of the product due to the quick and constant feedback at each phase of development. Each phase is listed below for reference:

- Plan requirements – Gather requirements of the product through communication with stakeholders.
- Develop product – development of the software based on the defined requirements
- Test software – Test products for errors and its ability to meet the user's acceptance criteria
- Delivery – Release of the product into production
- Incorporate feedback – Receive feedback from stakeholders to generate requirements for the next iteration of the product

Funding Requirements

The initial funding required to implement this social media analyzer is approximately $50,000. Much of the initial funding would be used on development costs and human resource-related costs. It will include any in-scope development time with the proposed project. If additional requirements are added during the agile development cycle, which is out of scope, additional funding will be required. The additional funding will be calculated based on the complexity of the new requirement and the development time needed to meet the new requirement.

Stakeholder Impact

The machine learning tweet analyzer would have a positive impact on 3Delta Consulting Inc. and its clientele by hiring employees that are compatible with the company culture and mission. It will allow 3Delta Consulting Inc. to have a positive work environment, and to build trust with their diverse clientele. In the long run, this will help cement relationships that will result in growth for both 3Delta Consulting Inc. and its clients.

Data Precautions

The data that will be used in the development of the machine learning tweet analyzer is publicly available data that does not involve or violate HIPPA, FERPA, or PCI DSS regulations. The data product, both while in development and production, will not contain any personally identifiable information (PII) that is not already publicly available, nor will any of the data be stored long term within any databases.

Developer's Expertise

The developers that will be assigned to this development team are recent graduates from Western Governors University. The Developers have some experience in Python, but not its machine learning libraries. However, they have previous training and work experience in web development, which will prove valuable in the development of the User Dashboard. They have also worked on many teams using the agile development process, which will prove valuable to the development of the product and company. There is no doubt that this team will be able to develop the twitter analyzer efficiently and timely.

Project Proposal

Problem statement

3Delta Consulting Inc. would like to update our hiring and employee retention procedures. Due to an unfortunate situation with a former employee, it has become necessary to analyze future applicant and current employee tweets for hate speech. The analyzer will use machine learning algorithms to identify any hate speech that an applicant or employee has engaged in through their tweets. Once analyzed this information will be viewable on a dashboard and will help the company to hire unbiased employees, and maintain a positive culture and relationships with clientele by rooting out prejudice and unethical behavior.

Customer Summary

The machine learning analyzer is meant to be an intuitive application that can be used by members of management, such as hiring managers and team leads, as well as the Human Resources (HR) Department. Due to the intuitive and user-friendly nature of the dashboard, no special skill set will be required to use the application. Members of management and the HR Department will be able to quickly verify if an applicant or employee has participated in hate speech on Twitter, and then act accordingly based on the summary given. They will be able to do this conveniently from their employee portals by using the link provided.

Existing System Analysis

Currently 3Delta Consulting Incorporated does not have any form of applicant or employee tweet analysis. Up until this point, the company has relied on employee referrals, recommendation letters, and human judgment when hiring and retaining our employees. Unfortunately, this process is vulnerable to human error and is ultimately insufficient. The machine learning analyzer reduces the risk of human error, and biases, and is much more time-efficient in providing results, than a manual search and analysis, would be.

Data

The data used to develop the Twitter Profile Analyzer is a dataset that is available on Kaggle.com (that data can be found using the following link: https://www.kaggle.com/datasets/mrmorj/hate-speech-and-offensive-language-dataset). The dataset contains the following columns:

- Index (Data Type: integer)
- Count (Data Type: integer)
- Hate Speech (Data Type: integer)
- Offensive Language (Data Type: integer)
- Neither (Data Type: integer)
- Class (Data Type: integer)

- Tweet (Data Type: string)

The machine learning tweet analyzer uses the class and tweet data that is in the dataset because the rest of the data is unimportant or unnecessary to the machine learning model. In the future, when the ML model is implemented into the 3Delta Consulting Inc.'s hiring process, the data set will be larger than the dataset used. It will lead to a more accurate analysis of tweets and, by extension, the Twitter profile due to additional data points. The data pulled from Twitter should be just the text from each tweet.

Project Methodology

The project will be developed using the Agile methodology. This will allow for constant communication from the development team to the stakeholders, and vice-versa, to ensure all requirements are met and any issues or alterations that occur during development are known to all parties. The agile methodology phases are as follows:

- Requirements Gathering –The development team communicates with the stakeholders to identify the requirements; the requirements can change due to the feedback received from the stakeholders.
- Develop – The software is developed to meet the requirements outlined in the previous phase.
- Test – The agile methodology promotes continuous testing, which is performed by both the development team and the Quality Assurance (QA) team. Our development team will be using unit testing to test the logic and check that all code is being covered, while our QA team will be using black-box testing to test the usability of the application.
- Deliver – The product will be delivered in three phases [Alpha, Beta, and General Audience (GA)]. During the Alpha phase, the product will be delivered to key stakeholders (HR managers). During the Beta phase, the product will be delivered to 20 percent of the target users for user usability. During GA, the product will be released to recruiters, hiring managers, and Human Resources Department.
- Feedback – Feedback creates new requirements and is generated from the stakeholders and is brought in in the next iteration of the product. If additional features are out of scope for the project, additional funding may be required.

In the agile methodology, many of these phases may overlap and can be automated. To speed up development, all testing by the development team will be fully automated. The team will also be using a Continuous Integration Continuous Deployment (CI/CD) pipeline. This will help not only with development in the future but also with quality testing.

Project Outcomes

The project deliverable should include any material that is created during the product development. The deliverables will include the schedule of the product development, the budget distribution of the project, documentation of any issues that arise during development, and development documents such as storyboards and flow charts.

The product deliverable will include the Twitter Profile Analyzer and the data visualization dashboard and the application programming interface (API) which will facilitate the communication between the analyzer and the dashboard. The source code and the dataset of the product will also be included in the deliverables.

Implementation Plan

The primary approach for implementation will use the Agile methodology. Initially, requirements and milestones will be produced through interactions among the development team and the stakeholders. The implementation rollout will have five phases, which are the Initial Prototype, Main Development and testing, Alpha Testing, Beta Testing, and General Audience Release to the stakeholders.

Upon beginning the development phase, the development team will work closely with the 3Delta Consulting Inc.'s hiring managers to continuously test and improve the prototype of the Twitter Profile Analyzer. Testing will be done by the development and the QA team throughout the development of the product to ensure the logic and functionality of the product have minimal errors and that all code is being tested. Upon the product passing quality assurance and meeting the requirements set forth, the product will go to Alpha Testing to test the usability of the dashboard and the information that is presented to them. After a week in Alpha testing, and no major issues being found, the product will be promoted to Beta Testing for testing by many of the hiring managers, and the HR Department. After another week with no major issues, the product will be considered ready for general usage. The phases of development are put forth as thus:

- Phase One – Development and Testing of a working prototype of the Tweet analyzer and dashboard
- Phase Two – Development and Testing of a functional model of the product. Unit Testing and Black box testing will be utilized for functionality and usability testing
- Phase Three – Acceptance Testing by key stakeholders, and review and implementation of feedback from stakeholders
- Phase Four –Beta Testing by key stakeholders and a selective few hiring managers
- Phase Five – Release to all hiring managers and the HR Department

The most important deliverable that will be delivered at the end of the Agile development cycle will include a fully functional Twitter Profile Analyzer with all documentation regarding errors and implementation. Other deliverables include the data visualization dashboard and other documents from the planning and design phase of the product development.

Evaluation Plan

Implementing the Agile methodology, the Development and QA teams will constantly test the product at all stages of development. Any errors and missing or incomplete requirements will be identified quickly and corrected before the deployment into a live production environment. This will be accomplished through a joint effort and communication of the development teams and the stakeholders.

The machine learning analyzer will be evaluated based on the accuracy of the identified tweets as hate speech, which can be subjective and dependent on the user. Initially, the Alpha Testers will be used to measure the accuracy of the analyzer, but the data sample size may be too small to accurately determine anything conclusively. During beta testing and production deployments, there will be a rating system that will be implemented for the managers to rate the accuracy, usability, and satisfaction of the analyzer. By collecting the user feedback an accurate measurement of success can be made.

Resources and Costs

Programming Environment

The data product will use Python 3.6.9 using Google's Colaboratory, Python data analysis, and machine learning libraries, along with several JavaScript libraries to develop the Twitter analyzer and dashboard.

The resources are compatible with both Mac and Windows computers, they are open-sourced and free of cost. The programming environment will effectively be $0.

Environment Costs

The product is an implementation of a tweet analyzer, Node.js API, and data visualization dashboard, which can be hosted on AWS. There would be a small associated cost of the hosting, while there would be no cost related to licensing. Effectively the environmental costs will be less than $100 during development and while in production per year.

Human Resource Requirements

The development of this product will involve a lot of development hours from our staff, which the initial quote of the cost is $49,800. Out-of-scope requirements can increase this cost. The time and cost of the development phases are broken down as follows:

- Phase One – 120 hours ($12,400)
- Phase Two – 240 hours ($24,800)
- Phase Three – 40 hours ($4,200)
- Phase Four – 40 hours ($4,200)
- Phase Five – 40 hours ($4,200)

Timeline & Milestones

The timeline and milestones for the proposed project are in the table below. The approximate time this project will take to complete is about one month and two weeks.

Milestones	Event	Start Date	End Date	Duration (in hours)	Dependencies (Tasks)	Resource Assigned
1	Project Requirement Meeting	3/21/2022	3/21/2022	8.0	NA	Development Team QA Team StakeHolders
2	Project Kickoff Meeting	3/22/2022	3/22/2022	8.0	1	Development Team QA Team StakeHolders
3	Product planning	3/23/2022	3/25/2022	24.0	2	Development Team, QA Team
Phase 1						
4	Product Design	3/28/2022	4/1/2022	40.0	3	Development Team, QA Team
5	Data Analysis	4/4/2022	4/8/2022	40.0	4	Development Team, QA Team
6	Deliver Prototype	4/11/2022	4/15/2022	40.0	5	Development Team QA Team StakeHolders
Phase 2						
6	Tweet Analyzer Development Including QA Testing	4/18/2022	4/29/2022	80.0	3, 4, 5	Development Team, QA Team
7	Dashboard Development Including Acceptance Testing	5/2/2022	5/13/2022	80.0	6	Development Team, QA Team
8	Incorparate Feedback into Product	5/16/2022	5/27/2022	80.0	6,7	Development Team QA Team StakeHolders
Phase 3						
9	Alpha Testing	5/30/2022	6/3/2022	40.0	6, 7, 8	Stakeholders
10	Improve product with feedback from Apha Testing	5/30/2022	6/3/2022	40.0	6, 7, 8	Development Team, QA Team
Phase 4						
11	Beta Testing	6/6/2022	6/10/2022	40.0	9, 10	Users Selected for Beta Testing
12	Improve product with feedback from Beta Testing	6/3/2022	6/10/2022	40.0	9, 10	Development Team, QA Team
Phase 5						
15	Product Release	6/11/2022	6/11/2022	40.0	1-12	Development Team QA Team StakeHolders
Total				600.0		

Post-Implementation Report

Project Purpose

The purpose of this project was to help assist hiring managers and human resource personnel for 3Delta Consulting Inc. to filter out applicants and employees participating in hate speech on Twitter. The goal was to ensure a safe work culture and relationship with both employees and clientele. This was accomplished through a Twitter profile analysis application. The application uses machine learning to

score tweets and has a data visualization dashboard to view the results. The use of this new tool reduced the risk of human biases and is considerably quicker than manual analysis.

Datasets

The dataset used to create the dashboard and Twitter profile analyzer application was retrieved from Kaggle.com (https://www.kaggle.com/datasets/mrmorj/hate-speech-and-offensive-language-dataset). The data is stored as comma-separated values (CSV) file. Table 1 is an example of the raw data that was included in the CSV file:

	count	hate_speech	offensive_language	neither	class	tweet
0	3	0	2	1	1	3 things i cant stand while dealing with a bit...
1	3	0	3	0	1	“@NoRapist: on my way to fuck ur bitch h...
2	3	0	3	0	1	RT @BittyTang: “@Weed_Cloudz: Us calling...
3	3	0	3	0	1	This is why people are lucky they only deal wi...
4	3	0	3	0	1	“@WavyGodd: I don't like a ghetto female...

Table 1: First 5 Rows of CSV Data

Data Product Code

The analysis of the data can be broken up into several key elements. Namely, loading, exploring, preparing the dataset, and building and training the model. These worked together to create a model that had a high accuracy rate of predicting whether the tweets had 'hate speech' or not.

First, we needed to load the data; this can be done in many ways. Since we had a CSV file, it was quick and easy to upload the file to Google Drive. Then we connected Colab to Google Drive to read the file directly from there. In Figure 1, you can see how this was implemented easily in a simple function.

```python
def get_data():
    data_path = "/content/drive/MyDrive/Computer Science Capstone - C964/labeled_data.csv.zip"
    df = pd.read_csv(data_path, index_col=0)
    df = df.sample(frac=1).reset_index(drop=True)
    return df
```

Figure 1: Importing Data

We then needed to explore the dataset, so we knew what kind of data it contained. Along with getting the facts and summary of the data, we produced several graphs that show the distribution of the data. This allowed us to see where our data had issues, and where its strengths were.

Cloudflower (CF) allowed voting on the tweets in their dataset. It could have as many as nine votes and a minimum of three votes. The data was separated into three categories: hate speech, offensive language, and neither.
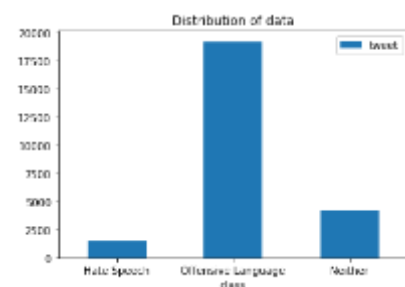

Figure 2: Pre-Equalization

In table 2, we can see that classifying hate speech can be biased, therefore, to reduce bias we decided that for the prototype we would reclassify all the tweets so that a single vote would up the severity. The severity of the tweets went from neither to offensive language, then finally to hate speech. This allowed us to create a greater balance between hate speech to neither. To avoid a type one error or a false positive, in production we would need a larger data set that is balanced in the types of tweets that we would like to detect.
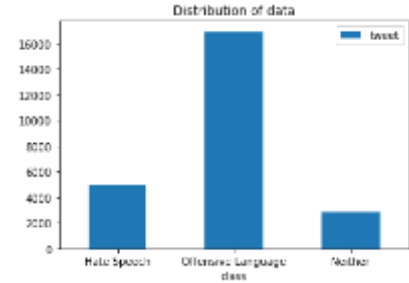

Figure 3: Reclassification of Data

	count	hate_speech	offensive_language	neither	class	tweet
0	3	1	2	0	1	@riwhey_000 I'm just kidding, I'm not a fuckin...
1	3	1	2	0	1	@mabathoana garbage rappers can't learn. it's ...
2	3	0	3	0	1	Im trying to take the word "bitch" outta my vo...
3	3	0	0	3	2	When all else fails, blast bare nips
4	3	0	0	3	2	RT @danwashburn: The Man Who Captured Muammar ...

Table 2: Classification Bias

Once we had our data in the two categories for the prototype we equalized the data, by sampling our hate speech category for the same amount of our neither tweet category. This would give us an equal number of hate speech to neither tweets, which is roughly 5,744 tweets.

Next, we split our data into two groups: the training dataset, and the test dataset. To do this we did an 80-20 split. This gave us 80% of our five-thousand tweets as our training data, and 20% of our data as test data.
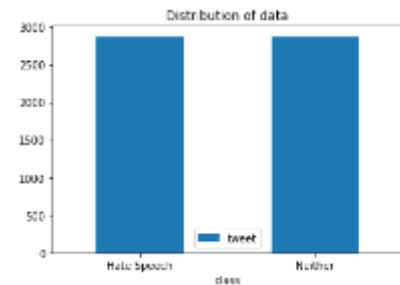

Figure 4: Equalization of Data

Finally, we created and trained our model. We started with our encoder. This layer would normalize data by converting all text to lowercase and stripping any punctuation. It then splits the sentence or sentences into substrings, typically words, and recombines them into tokens, usually n-grams. This layer then associates a unique integer value with each token and transforms each sentence using this index into one of the following two: a vector of integers, or a dense float vector. This all happens behind the scenes. Next, we retrieved and sampled our encoder to ensure all was working before building and training our model.

Once we had a working encoder, we started working on building the model. This was accomplished using one embedding layer, two bi-directional Long Short-Term Memory (LSTM) layers, a dense layer, a dropout layer, and a second dense layer. To make sure our model would be able to take the raw text in, we verified which layers in our model supported masking (the integer in place of the word). Next, we ran several texts through the model to verify that it was working.

After compiling and getting the summary of our model we were ready to train the model with our training data, and then test it. To begin with, we trained our data with 10 epochs or iterations. This resulted in an accuracy of about 99.71%. To verify that everything was working we ran several known strings through the model and tested how these were scored.

Hypothesis Verification

We hypothesized that through this tool we would be able to mitigate some business client damage and foster good faith while increasing our brand recognition by hiring employees that would help improve our company culture. We believe that this tool will become a great resource that will allow us to have positive relationships with our clients and will help increase their trust in us and our employees. Additionally, initial testing of the Twitter Hate Speech Detection shows that it can accurately predict hate speech.

Effective Visualizations and Reporting

By using graphs and charts throughout the project, we could visualize the data as it traveled down our pipeline to our model. By using the charts, we were able to visualize which columns we needed and which we did not, allowing us to remove unnecessary data that could confuse the model. We were then able to analyze the data through the graphs to see that the data was skewed to having little hate speech, although some of the offensive language tweets had votes to be considered hate speech. The charts were also able to help us show what amount of data we had and where we could adjust the vocabulary size that the model can learn. Also, through the printout of the summary of the model, we were able to see the shape of each layer and what it takes to create and train our model. The final graphs were able to show us how our data was performing and what kind of loss and accuracy it was obtaining.

Accuracy Analysis

Through the accuracy analysis and the charts generated we were able to see the accuracy percentage and the loss percentage of the data being run through the model. As discussed earlier, the accuracy of the Twitter analyzer will not be conclusive until it has been implemented within 3Delta Consulting Inc.'s hiring process, and user feedback from managers and HR is collected. This will be done by sending out periodic surveys to the teams, allowing us to know where we can improve through a rating system of 1-5. We will also implement a voting system on tweets that managers and HR believe were misclassified, to continue to learn and improve our model. This will allow us to monitor the health and accuracy of the Twitter profile analyzer. However, initial testing shows the machine learning model is highly accurate because it has been limited by the small dataset that was used for training.

Application Testing

The prototype was tested using integration testing and unit testing. Continuous testing was done per the agile methodology. The teams used the unit testing to validate that 90% or greater of code was being tested. Any portion of the code that was not needed and not being tested was removed. Once the application was complete, it was integrated with a mock-up of the Twitter API to be used by the Alpha testers. Once the duration of the Alpha testing phase was completed, then the application was fully integrated into the Twitter API, and it was released to the Beta group. Once all testing phases were complete, the product was released to all hiring managers, HR Department employees, and executive teams.

Application Files

The application is hosted on Google's Colab, which can be found at https://colab.research.google.com/drive/1pME8Q7G2mF6Jy_0wjwVt2PPD5Rt3jK7u?usp=sharing. The Capstone project Colab includes the following files:

- WGU Capstone.ipynb is a Google Colab Jupyter Notebook file of the Twitter tweet analyzer
- labled_data.csv is the dataset used to train and test the machine learning algorithm

User's Guide

1. Open Google Colab by navigating to colab.research.google.com

2. Using your Google credentials login to Google Colab
3. Upload the provided labeled_data.csv.zip
4. Select Runtime, then Run All
5. Once finished scroll to the bottom of the page
6. At the input marked "tweet" change this to any kind of sentence or tweet.
7. Rerun this code cell
8. The model should show a prediction: "Hate Speech" or "Neither"

Summation of Learning Experience

The Capstone Project has pushed me to the limits of my knowledge and tested my abilities in programming and machine learning. My previous experience in other WGU classes such as Software Engineering and Data Structures and Algorithms II has aided in the software development life cycle aspect of this project. As I was new to data analytics, I learned a lot of new concepts. I had to dig into and find answers to many questions and concepts, such as machine learning algorithms, their implementation, and how the many different frameworks work together to create a completed project. When I was stuck on a problem or needed help, I was able to reach out to Mark Denchy, Jim Ashe, and to my advisor Tammy Migliore to receive help and feedback. With their experience and expertise in different areas, they were able to help me complete my project. Without their assistance, I would have taken much longer and struggled far more to complete it. Their encouragement throughout the process of getting my degree was invaluable. I particularly appreciated Tammy's continuous encouragement along with Mark Denchy's time. Also, when a meeting with Mark or Jim was not easily created, I scoured the internet and school resources for solutions to the errors that I was getting. The culmination of all these resources provided the necessary environment for me to succeed in getting my degree.