

Abstract

Purpose: The sole intent of this research paper is to **study the attitudes and sentiment of students towards various services provided by the University of North Texas**. The feedback is collected through survey forms, which consisted of 7 open ended and 3 Likert Scale questions, among students pursuing **master's in information systems and technologies (MS-IST)**. Students have been asked to provide their reviews on Academic and Career Advising, Financial Aid, Dining, infrastructure, Library, Transportation and Recreational services and also about their confidence of getting a job after graduation and opinion on career fairs and how helpful is the course curriculum for getting a job and then the responses were cleaned and processed to know the emotion of the students on University of North Texas (UNT).

Through this research paper we were able to **determine the overall sentiment of students on UNT** student reviews.

Introduction

Abraham Lincoln once said, "With public sentiment, nothing can fail; without it, nothing can succeed" [7]. The Gallup report clearly shows that 51% of students in the USA are dissatisfied with their school experience [8]. This dissatisfaction has resulted in an increase in school dropouts, excess stress rate in youth, loneliness, crime, etc. Unless and until we understand the root cause by knowing it from the students themselves it would not be possible to improve the student's school experience. Students in the age of bombardment of jet age expectations are getting deprived of the attention that they should receive, from parents, teachers, etc., and are not able to open up to share their thoughts on their school life. So, to improve the educational services, to identify the learning gap and to enhance student engagement one needs to know the voice of students, and this can be done using Sentiment Analysis.

What is Sentiment Analysis??

Sentiment Analysis is a process through which one can identify whether the students (or customers) have a Positive, Negative or Neutral opinion towards the services provided by an institute. As part of the Sentiment Analysis process, one needs to collect the reviews or feedback from students and then analyze the responses to identify the attitude and voice of students.

Sentiment Analysis is a computational calculation that helps us systematically identify, extract, study, and categorize a particular text as positive, negative, or neutral. It is generally performed on the responses collected using rule-based approaches, Machine learning algorithms/ Natural Language Processing techniques, and in some cases a hybrid combination of both. Sentiment Analysis is also used by many companies to know the attitude of their existing and potential customers by analyzing and evaluating their social media posts, online reviews, etc.

The first step in the sentiment analysis process is **Data collection** which involves obtaining incoming data from different sources and in many formats. It is followed by **Data cleaning** where we remove or replace erroneous and null value data with appropriate data. Since the incoming data is in an unstructured format, we must transform it into a structured one for analyzing it. There

are different levels at which analysis can be done based on requirements i.e., Document level, sentence level, and aspect level. **The sentiment analysis in this research study was performed at the sentence level which is very popular to distinguish subjective and expressive statements from factual sentences for a more granular analysis of the text and to provide insights.**

After this, we analyzed the cleaned data of responses to determine the sentiment using a Machine Learning model called **Naïve Bayes Classifier** which is a probabilistic algorithm that uses the frequency of words in a document to predict the sentiment of the feedback text to analyze and understand the sentiments of the text feedback provided by graduate students towards various aspects of the University of North Texas.

Literature Review

In this section, the work demonstrated by different experts on several techniques for performing Sentiment Analysis on the customer/students' reviews. Normally, the study of sentiment analysis is carried out in various ways, but the most popular way of analyzing is using a well-trained machine learning algorithm.

In paper [1] author Thanveer Shaik et al. highlight the significance of opinion mining and sentiment analysis in educational settings, where it can be used to extract insights from student feedback, online forums, and social media platforms and their applications in various domains. Also, how it helps in improving student engagement, predicting student performance, and identifying areas for curriculum improvement.

In paper [2], the author discusses the advantage of having qualitative questions open-ended questions over Likert questions in the design of the survey design. Also discusses the need and possible ways of streamlining and automating the text analysis process.

In paper [3], the author uses the Module Evaluation Survey form that was conducted in the college in which they included some qualitative open-ended questions along with demographics and Likert Type questions. In the research study, they perform sentiment analysis using various Machine Learning Algorithms NN, K-NN, NB, and SVM using an open-source tool "*Rapid Miner*". The results of this research paper show that among all, the Naïve Bayes algorithm gave the best accuracy and recall of 97.07%.

In paper [4], the authors explain that Lexicon-based and machine learning-based methods make up the two primary types of sentiment analysis. Lexicon-based techniques determine the emotion of words and phrases by using a collection of lexicons that have been manually developed. Machine learning-based techniques employ machine learning algorithms to extract the sentiment of words and phrases from a training dataset.

In paper [5], the authors discuss about Opinion Mining (OM) which refers to the task of extracting opinions from text and Information Fusion (IF) which is the process of gathering information from multiple sources to have more accurate dataset. One of the biggest challenges of OM is the ambiguity of natural language. Another problem is the variety of text formats. OM

systems must be able to handle a variety of text formats, such as news articles, blog posts, and social media posts.

In paper [6] the authors went on to explore the efficiency of various machine learning algorithms in sentiment analysis on 2000 movie reviews in Twitter feed. They found that Naïve Bayes was most accurate with 89 percentage precision as compared to SVM which was 88 and Random forests with a precision value of 85%. Small data sets and the use of just one parameter (accuracy) to assess the effectiveness of the algorithms were two drawbacks of the study. By utilizing larger datasets and measuring the algorithm's effectiveness using more metrics, future work can get beyond these restrictions.

Methodology

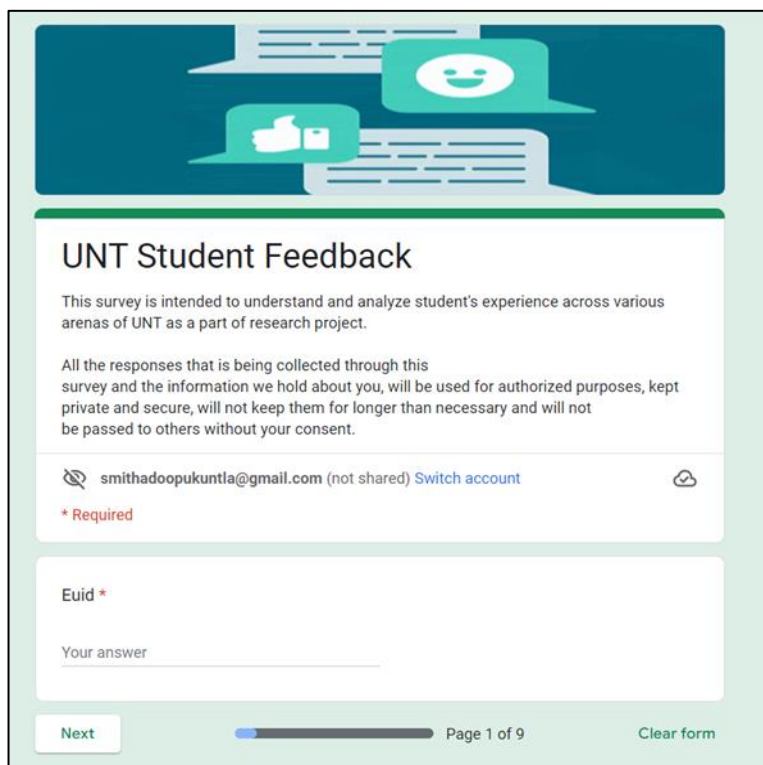
The major phases of Sentiment Analysis we have gone through for this research study consist of the following major phases.

- i. Survey.
- ii. Data Collection
- iii. Data Processing and cleaning
- iv. Determining the Sentiment

i. Survey

For collecting data, we used Google Forms which is an open-source survey administration software and can be used easily on mobile, laptops, tablets, etc.

The Google form asks students to provide their EUID as a unique identifier and consisted of 10 questions out of which **7 questions were open-ended** questions about advising, administration, parking, transportation services, offline and online course structures, and recreational activities. The other set of **3 questions were Likert scale questions** about the job opportunities after graduation, curriculum, and career fairs where "1" refers to poor, "3" average, and "5" excellent. Through the survey, we collected both quantitative data through Likert scale questions and qualitative data using closed-ended questions. Thus, making the study of mixed nature.



The screenshot shows a Google Form titled "UNT Student Feedback". The form has a header with a blue background and white icons of a laptop, a speech bubble with a smiley face, and a thumbs up. The main content area is white with a green border. It includes a title "UNT Student Feedback", a description "This survey is intended to understand and analyze student's experience across various arenas of UNT as a part of research project.", a privacy notice "All the responses that is being collected through this survey and the information we hold about you, will be used for authorized purposes, kept private and secure, will not keep them for longer than necessary and will not be passed to others without your consent.", a user identification field "Euid *" with a red asterisk indicating it is required, and a "Next" button. The footer shows "Page 1 of 9" and a "Clear form" link.

ii. Data Collection

The online survey form was circulated among all the students at UNT who were pursuing masters, especially in Information Systems and Technologies. Data were collected from 15 March 2023 to 30 March 2023. We forwarded the survey form link through various platforms like Canvas, WhatsApp, Group me, etc. requesting students to provide their feedback. We were able to collect data from 145 students but among it, due to the inappropriate data, we could consider only 140 students' data for this research study. The response rate of students was 13.5% which is very low because students were irresponsible and hesitant to take up the survey and were very passive to participate in the survey.

iii. Data Processing and cleaning

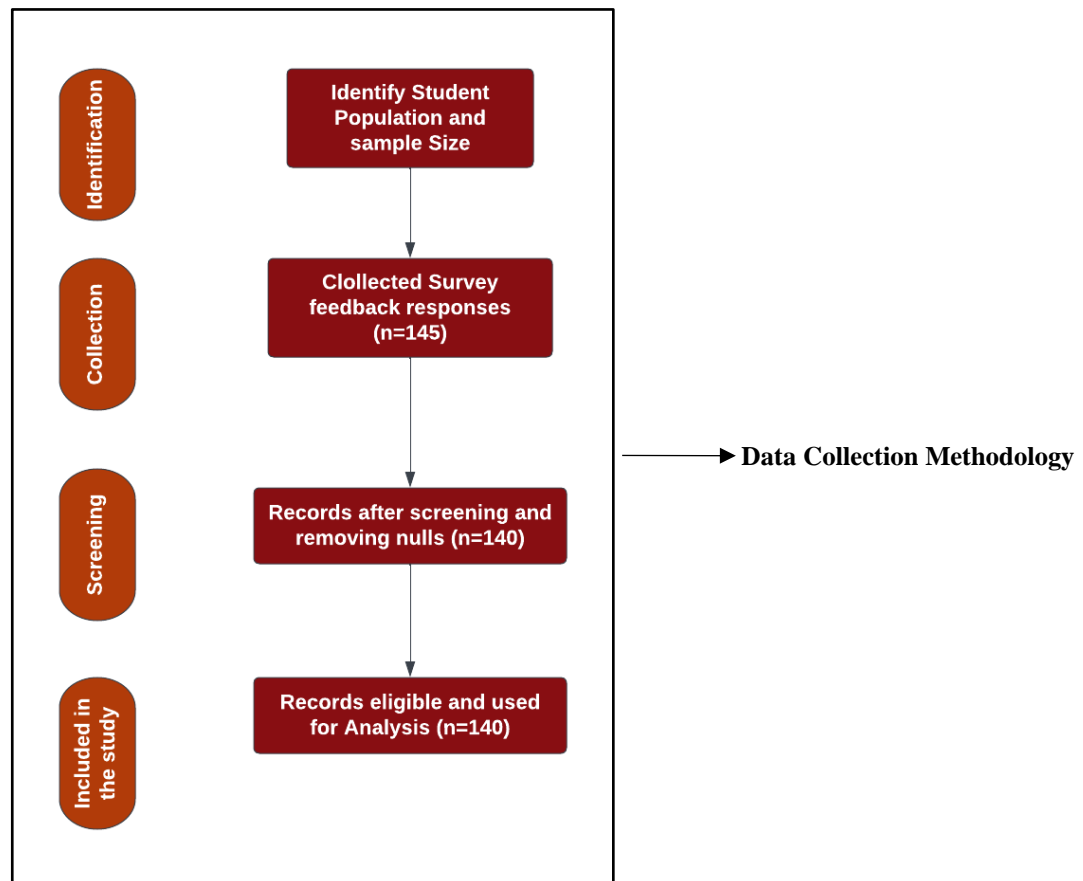
As the data is unstructured, we need to clean it up and make it ready before using it for analysis. Out of three, document level, sentence level, and aspect level this research deals with sentiment analysis at a sentence level granularity so we need to remove all the redundant data and keep only those strings that help in finding the emotion.

Stemming - a process of removing redundant words, like suffixes or prefixes. The resulting stem may not be an actual word, but it is often useful for text analysis and information retrieval.

Tokenization - a process of breaking a text into individual words or tokens. This is usually done by separating words based on whitespace or punctuation.

Part of speech tagging – it is used to determine whether a word is a verb, noun, or adjective that will be helpful to detect the emotion.

Bag of Words - a text representation technique that is mostly used for feature extraction. It represents each sentence as a vector of the occurrence frequencies of all the words in the corpus, disregarding the order and structure of the words. In other words, BoW is a simple way of converting text into numerical data that can be processed by machine learning algorithms.



Parsing - the process of analyzing the syntactic structure of a sentence and identifying the relationships between words. This can be done using techniques such as dependency parsing or constituency parsing.

Machine Learning Model Selection:

Now we are ready to use it for analysis. At this point selecting the most suitable Machine learning algorithm was a challenging task. In the process of selection of a suitable Machine Learning Model one needs to consider factors such as Data Size and Quality, Data Type (Text or Numeric), Type of Sentiment Analysis, Performance Metrics, Interpretability, Handling of Noise and Outliers.

According to the scikit learner if the data set is less than 100,000 and the data is in text then the most suitable Machine Learning Algorithm is **Naïve Bayes**, and it is because of the following reasons.

1. Works Well with Small Data comparative to other algorithms
2. Requires Less Training Data
3. More Efficiency
4. Higher Accuracy

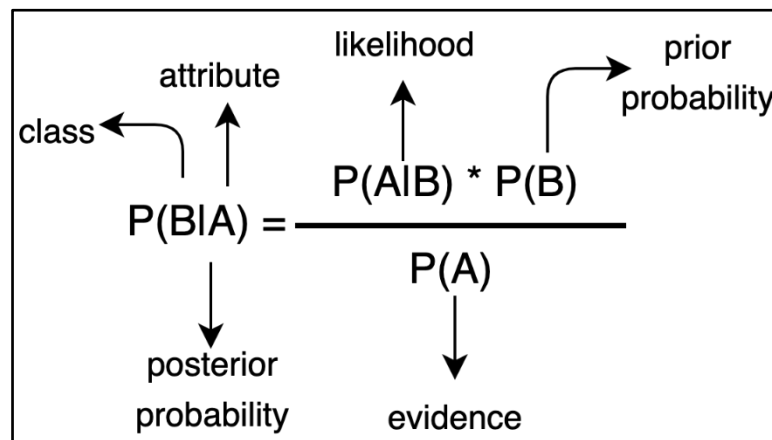
Student Feedback data set after cleaning and processing.

3Q - Scholarships	4Q - Recreational	5Q-Dining	ONLINE/Offline Pref	Reason
yesireallyhappywithfinancialaid	yes	goodaverage	online	onlinecoursesomecourseshaverecordingsessionwecanlistenandlearnmanythingsimple
good	good	yes	online	easytoacquireknowledge
poor	yes	average	online	itsjustmyopinion
good	good	better	inperson	good
poor	good	average	online	donno
ifeelitsgood	yes	expensive	online	itsmoreconvenient
asaninternationalstudentiamterriblydissat	verygood	good	online	onlineclassareeasytoaccess
forinternationalstudentsitisbitexpensive	ienjoyoncampusfacilitie	foodisworththepri	inperson	inpersonclassisgood
yes	no	midrange	online	opinionmaydifferbasedoncoursewetake
bad	good	good	inperson	ipreferinpersonbecausewecaninteractwithlotsofpeople
good	good	bad	inperson	moreexposuretopeopleandprofessor
bad	good	good	inperson	ipreferinpersonclasseswithonlineexams
yes	no	average	online	onlineisbest
itsveryhelpfultome	yes	goodfoodinaverag	online	onlineclassesasierbetterthaninperson
noteasyforeveryone	yes	average	inperson	itismoreinteractive
p	better	yes	online	yes
asaninternationalstudentifeelthatimnotsa	inmyopinionthereareple	thefoodoncampus	inperson	theinersonclassesprovidemoreflexibilitytointeractwiththeprofessorsandclarifyallthe
good	lesscardiomachinesandr	expensive	inperson	ididallonlineforyearsnowimdoingallonlineitislessstressfulmoreengagingandilearnmore
somewhat	good	average	inperson	causeitsmoreengaging
dissatisfied	good	average	online	flexibilityinschedule
normallysatisfied	lotofactivitiesbutofotenp	average	inperson	interactionwouldbemoreandeffectiveincaseofofflineclasses


```
import pandas as pd
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, confusion_matrix
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
from sklearn.model_selection import train_test_split
```

The algorithm works as follows:

1. Given a set of training data with class labels and feature values, calculate the prior probability



of each class label. This is simply the proportion of the training data that belongs to each class.

2. For each feature value in the training data, calculate the conditional probability of that value given each class label. It is carried out by counting the no. of times each feature value appears in each class and then normalizing by the total no. of feature values for that class.
3. To classify a new input instance, calculate the posterior probability of each class label given the input features using Bayes' theorem:
4. Choose the class label with the highest posterior probability as the predicted class for the input instance.

Naive Bayes on the other hand is a simpler model that relies on basic statistical principles and has fewer parameters to learn. They are less prone to overfitting and can be trained effectively on smaller datasets.

Following are the major libraries that are used for text cleaning and analysis:

1. **Scikit-learn** (also known as sklearn) - used for classification, regression, clustering, and dimensionality reduction.

Scikit-learn packages.

- **CountVectorizer** - used for feature extraction and creating the document-term matrix. It creates a matrix representing the frequency of each word in a document.
- **MultinomialNB** - a classification algorithm based on Bayes' theorem, which is used to classify text data in this case. It is implemented in the sklearn library, making it easy to use in Python.

- **train_test_split** is a function from sklearn.model selection that is used for breaking the dataset into training and testing sets. This function helps in gauging the performance of an ML model on data which is not been seen before.
- The **accuracy score, confusion matrix, and classification report** functions are used for evaluating the performance of the classifier. These functions are also provided by sklearn.

2. Numpy - used for numerical computing. It provides support for arrays and matrices, which allow for efficient computations on large sets of numerical data.

3. Pandas - used for data manipulation and analysis. It provides data structures such as DataFrame and Series that make it easy to work with structured data.

4. Nltk -The Natural Language Toolkit (NLTK) is used for natural language processing. It provides tools for tasks such as tokenization, stemming, and sentiment analysis.

5. Re-module provides support for regular expressions in Python. Regular expressions are a powerful way to search and manipulate text.

6. Textblob - used to process textual data. It offers resources for activities including noun phrase extraction, sentiment analysis, and part-of-speech tagging.

7. Wordcloud - used to create word clouds, which are visual representations of text data where the size of each word is proportional to its frequency in the text.

8. Seaborn - used for data visualization. It provides a high-level interface for creating statistical graphics.

9. Matplotlib – needed to create static, animated, and interactive visualizations that help in understanding the trends and patterns.

10. Cufflinks - connects the Pandas data frame with Plotly enabling users to create interactive visualizations directly from Pandas.

11. Plotly: creating interactive and customizable visualizations. It provides a range of chart types, including scatter plots, line charts, and heat maps.


```

import numpy as np
import pandas as pd
import nltk
import re
from textblob import TextBlob
from wordcloud import WordCloud
import seaborn as sns
import matplotlib.pyplot as plt
import cufflinks as cf
%matplotlib inline
from plotly.offline import init_notebook_mode, iplot
init_notebook_mode(connected=True)
cf.go_offline();
import plotly.graph_objs as go
from plotly.subplots import make_subplots

```

Libraries used.

iv. Determine the sentiment.

Now we need to train the Naïve Bayes classifier by splitting the dataset 80 % for training the algorithm and 20% for testing and it follows as below.

1. The dataset is given as input and then we Split the dataset into training and testing sets.
2. Create a CountVectorizer object to transform the reviews into numerical features.
3. Transform the training and testing sets.
4. Create a Multinomial Naive Bayes classifier.
5. Train the classifier on the training features and sentiments.
6. Use the classifier to predict the sentiments of the testing features.

Here, the predicted sentiment will always be the one with the greatest probability.

```

Review 1: I think its quite good as the support services are effective
True sentiment: Positive
Predicted sentiment: Positive
Probability of negative sentiment: 0.0034
Probability of positive sentiment: 0.0009
=====
Review 2: Super
True sentiment: Positive
Predicted sentiment: Positive
Probability of negative sentiment: 0.1133
Probability of positive sentiment: 0.1581
=====
Review 3: In my opinion the academic advising service is good
True sentiment: Positive
Predicted sentiment: Positive
Probability of negative sentiment: 0.0281
Probability of positive sentiment: 0.0123
=====

```

The accuracy of the algorithm is 82% and the confusion matrix we obtained are as follows.

```
=====
Accuracy: 0.8214285714285714
Confusion matrix:
[[ 6  1  1]
 [ 1  1  1]
 [ 1  0 16]]
Classification report:
              precision    recall  f1-score   support

   Negative       0.75        0.75        0.75         8
    Neutral       0.50        0.33        0.40         3
    Positive       0.89        0.94        0.91        17

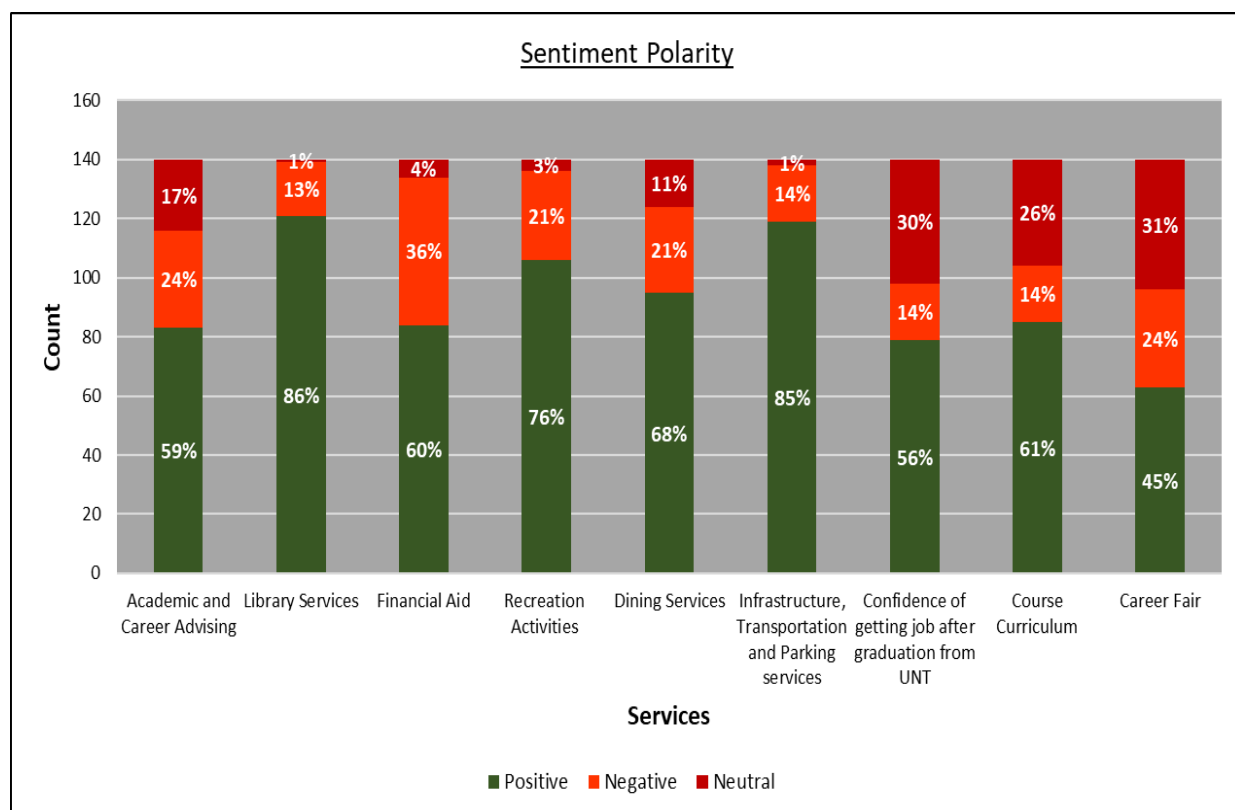
 accuracy              0.82        28
 macro avg           0.71        0.67        0.69        28
 weighted avg        0.81        0.82        0.81        28
```

Now the model is trained, and it can be used to determine the polarity of other reviews easily.

```
Sentiment counts:
 Positive      119
 Negative      18
 Neutral        3
dtype: int64
```

Sentiment Determined Using Trained Naïve Bayes Algorithm

Results



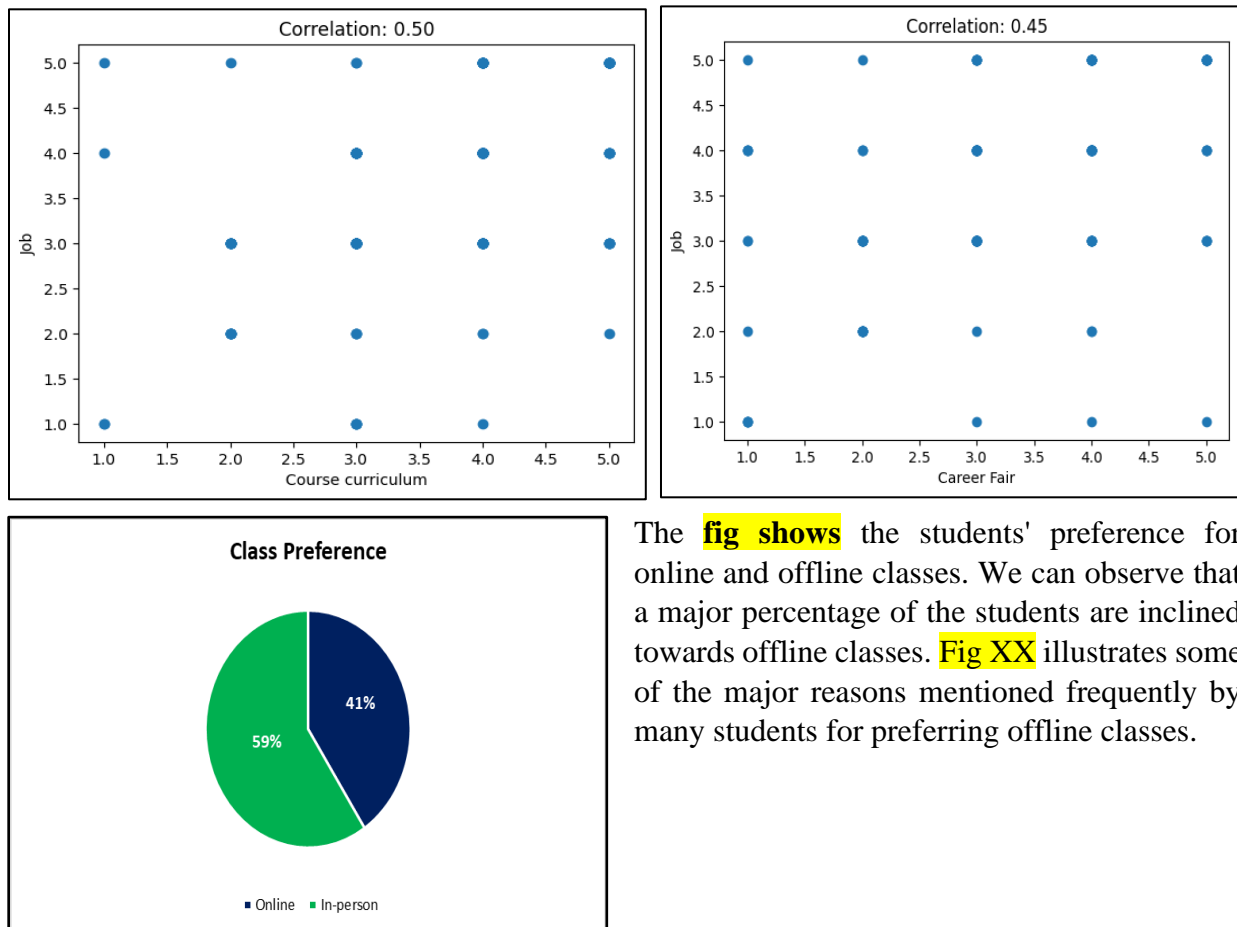
Based on the analysis performed using the naïve bayes classifier on the feedback received from the students on open ended questions for academic advising and career counseling, Library Services, Financial Aid, Recreational Activities, Dining services, Infrastructure, Transportation and Likert Questions Career fairs, and Curriculum we got the following results.

The above stacked chart illustrates the sentiment of students (Positive, Negative, Neutral) at the University of North Texas. Among all the divisions students showed highly positive responses on Infrastructure, Transportation, and Parking Services. The Library services also received a pretty good opinion from students. We can also make out that students are highly disappointed with the career fairs and confidence in getting a job after graduation from UNT and then with the Academic and Career Advising.

Students were more neutral towards Financial Aid and some of the major reasons that were mentioned frequently by many students are highlighted in the below word cloud. It clearly shows that students are pointing more toward communications, procedures, evaluations, and lack of proper communication. These word clouds are also a good source to identify ideas and improvements, like here students are telling about the communication gap, and lack of clarity regarding deadlines and criteria and are suggesting a better website to overcome.

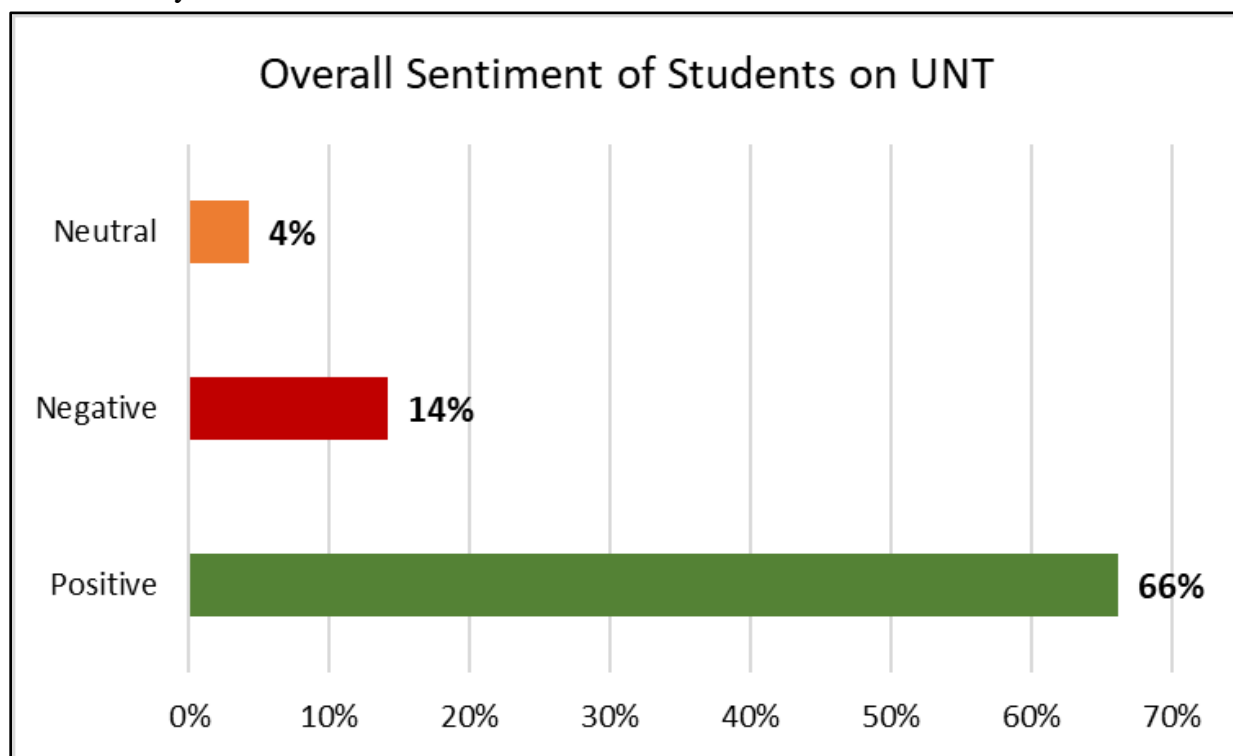
Career Fair	
Mean	3.642857143
Standard Error	0.085465552
Median	4
Mode	4
Standard Deviation	1.01124205
Sample Variance	1.022610483
Kurtosis	-0.14611365
Skewness	-0.545754485
Range	4
Minimum	1
Maximum	5
Sum	510
Count	140
Largest(1)	5
Smallest(1)	1
Confidence Level(95.0%)	0.168980586

The below figure (XX) shows the correlation between students feeling confident about getting a job based on their course curriculum and found that about 50 percent of students are appreciative stating that the courses were helpful, and they are assured of getting a job after graduation. At the same time figure (XX), we can observe that nearly the same number of students also said career fairs have nearly equal influence on their chance of getting a job.



The fig shows the students' preference for online and offline classes. We can observe that a major percentage of the students are inclined towards offline classes. Fig XX illustrates some of the major reasons mentioned frequently by many students for preferring offline classes.

We analyzed the overall sentiment of students on services provided by the University of North Texas. From the fig, it is seen that the students had positive feedback on the Services provided by the University of North Texas.



Recommendations

- Academic and career advising needs to be more active during the beginning of semester as it takes a lot of time to book an appointment with them and can be one of the main reasons for dissatisfaction.
- Financial aid has room for improvement wherein more scholarship awards can be introduced to support students financially by communicating students about the deadlines and qualifications to avail for the aid.
- Since the Covid has ended and already 59 percent of student population are showing their interest in in-person mode of classes to further promote it the class strength can be reduced, and class timing choices can be increased.
- Since, students say that their chance of getting a job heavily depends on their course curriculum and career fairs, so courses should be evaluated on regular basis if they are meeting the market requirements and also students should be prepared to best utilize the career fairs. Since Career fairs have relatively greater dissatisfaction level, more work needs to be done on this respect wherein we can have more frequent career fairs with better advertising in advance so that students can prepare in advance for them.

Challenges

1. **Selection of the Machine Learning Model:** **Selection of the Machine Learning Model:** Selecting a best suitable algorithm is a challenge [9]. There are many machine learning models available that can be used for performing sentiment analysis, but the major challenges was to select a particular model such that we avoid data overfitting and underfitting so for that we have studied various papers and observed the models used for different kinds of datasets and then compared our dataset and we chose the model accordingly.
2. **Survey Fatigue:** Students were reluctant to take the survey and the main challenge was to keep the students active and make them fill in the complete survey form.
3. **The ambiguity of language:** The feedback contained ambiguous language including jargon, improper English, and metaphors which made it difficult for the algorithm to understand and classify it.
4. **Unstructured data:** The data was not organized and was in different formats and lengths containing numbers and just a word at a time too long paragraphs in others.
5. **Biases in data collection:** Since this is a major challenge while attempting to collect data by oneself, we tried to mitigate this by asking random students from MS-IST department across different aspects of UNT services.
6. **Lack of domain knowledge:** It's observed that students, due to their lack of knowledge in various UNT services, are giving remarks which may not be exactly true and thus represent a wrong outlook of the overall student population.
7. **Less quantitative data:** Since the data was very limited and we couldn't get enough responses from students, analyzing data, finding correlations and regression was difficult.

Conclusion

This research study is dedicated to performing sentiment analysis and interpreting the student's emotions based on the feedback collected through the survey of various divisions of the University of North Texas.

The study gives an insight that the University of North Texas needs to improve a lot on career fairs, improving course curriculum such that they built confidence in students in getting a job after graduation and, they need to improve financial aid services.

On and overall **students have positive opinion** on major of the services provided by **University of North Texas**.

References

1. Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2023). Sentiment analysis and opinion mining on educational data: A survey. Natural Language Processing Journal, 2, 100003. <https://doi.org/10.1016/j.nlp.2022.100003>
2. Gronberg, N., Knutas, A., Hynninen, T., & Hujala, M. (2021). Palaute: An online text mining tool for analyzing written student course feedback. IEEE Access, 9, 134518-134529. <https://doi.org/10.1109/access.2021.3116425>

3. Dhanalakshmi, V., Bino, D., & Saravanan, A. M. (2016). Opinion mining from student feedback data using supervised learning algorithms. 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC). <https://doi.org/10.1109/icbdsc.2016.7460390>
4. Rajalakshmi, S., Asha, S., & Pazhaniraja, N. (2017). A comprehensive survey on sentiment analysis. 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN). <https://doi.org/10.1109/icscn.2017.8085673>
5. Balazs, J. A., & Velásquez, J. D. (2016, January). Opinion Mining and Information Fusion: A survey. *Information Fusion*, 27, 95–110. <https://doi.org/10.1016/j.inffus.2015.06.002>
6. Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of anti-LGBT campaign in Indonesia using naïve Bayes, decision tree, and Random Forest algorithm. *Procedia Computer Science*, 161, 765–772. <https://doi.org/10.1016/j.procs.2019.11.181>
7. Speech of hon. Abraham Lincoln at Ottawa [Concluded]. (n.d.). Northern Illinois University Digital Library. <https://digital.lib.niu.edu/islandora/object/niu-lincoln%3A38360>
8. Gallup. (2022, September 1). Americans' satisfaction with K-12 education on low side. Gallup.com. <https://news.gallup.com/poll/399731/americans-satisfaction-education-low-side.aspx>
9. In Lee a , Yong Jae Shin b. Machine learning for enterprises: “Applications, algorithm selection, and challenges” *Business Horizons* Volume 63, Issue 2, Pages 157-170. <https://www.sciencedirect.com/science/article/pii/S0007681319301521>