Employee Attrition & Sentimental Analysis

Made by: Abhishek Mishra

What is Employee Attrition?

- Putting simply, employee attrition is the reduction of staff by voluntary or involuntary reasons. These can be through natural means like retirement, or it can be through resignation, termination of contract, or when a company decides to make a position redundant.

Objectives

- This work aims in developing a Decision Support System in employee attrition detection that uses the data mining technique having best accuracy and performance among Support Vector Machine, KNN, and Random Forest etc.
- By using several HR employee data management system parameters such as satisfaction level, last evaluation etc.
- And finally, deploying the best model API.

Data structure

NO.	ATTRIBUTE	DESCRIPTION
1	Satisfaction Level	Employee is satisfied or not with his or her work (values range from 0 to 1). 0 stands for not satisfied and 1 stand for satisfied.
2	Last Evaluation	Last rating of employee value in the columns (values range from 0 to 1).
3	Number of Projects	Number of project employee is working on (2, 3, 4, 5, 6, 7 projects).
4	Average Monthly Hours	Average monthly hour spent
5	Time spent at the Company	Number of years spent in the company (2, 3, 4, 5, 6, 7, 8, 10 years).
6	Whether they have had a work accident	During working did employee have an accident 0 - no and 1 for yes.
7	Whether they have had a promotion in the last 5 years	promotion in last 5 years 0 - no and 1 for yes
8	Department	Employee working in which departments (sales, accounting, technical, support, IT, product_mng, marketing, HR, R&D, management).

Techniques Used

- **Random Forest**
- **Decision Tree**
- **KNN**
- **SVM:**

Outcomes

Technique	Accuracy
SVM	95.54%
KNN	96.0%
Decision Tree	92.23%
Random Forest	98.9%

Conclusion

- Through my analysis, managing the level of satisfaction is the key to keep employees with the firm.
- This is especially important for employees who have been around for more than 3 years. Other than that the employee evaluation and number of projects should also be monitored.
- This firm's HR Head would do well to craft programs to keep tabs and these metrics so as to have a successful talent retention policy.

Recommendation

- At, then end I would like the firm to adopt my ML API, as it will not only allow them to save cost due untimely quitting of job by an employee. But also able to retain them. And further this can be share with others companies or subsidiaries and prove to be commercial success.

Sentimental Analysis

- Definition

Sentiment Analysis is the process of 'computationally' determining whether a piece of writing is positive, negative or neutral. It's also known as opinion mining, deriving the opinion or attitude of a speaker.

OBJECTIVEs

- This work aims at developing a system which tells us about the sentiment prevailing for the particular industries, product etc.
- In this paper, I have used the random forest as the data mining technique and for visualization, I have applied “Word cloud”.
- The dataset deals with the sentiments of people for US-based airlines industry.

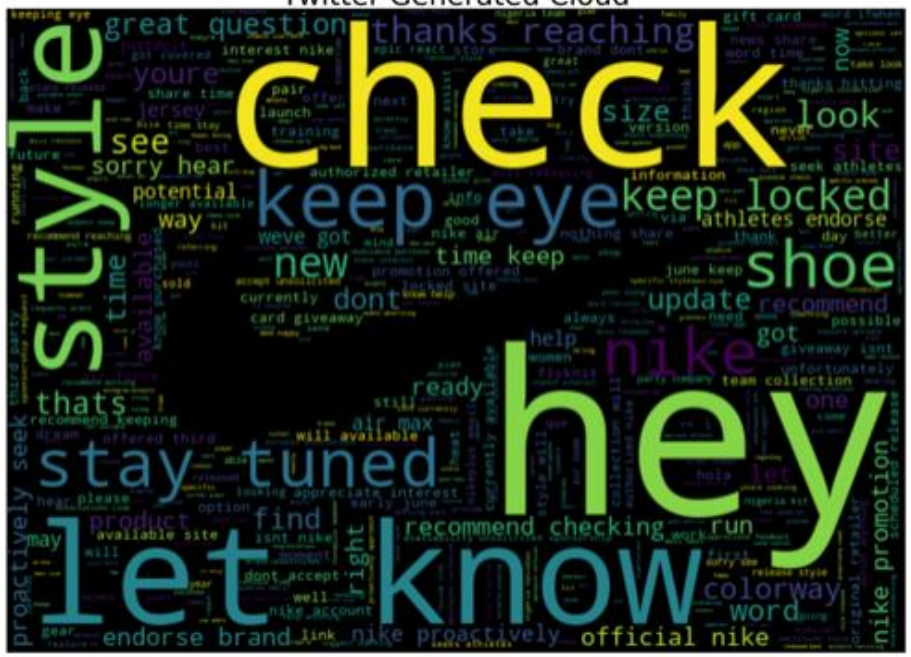
Technique Used

- Word Cloud: A tag cloud is a novelty visual representation of text data, typically used to depict keyword metadata on websites, or to visualize free form text. Tags are usually single words, and the importance of each tag is shown with font size or color.

Word Clouds

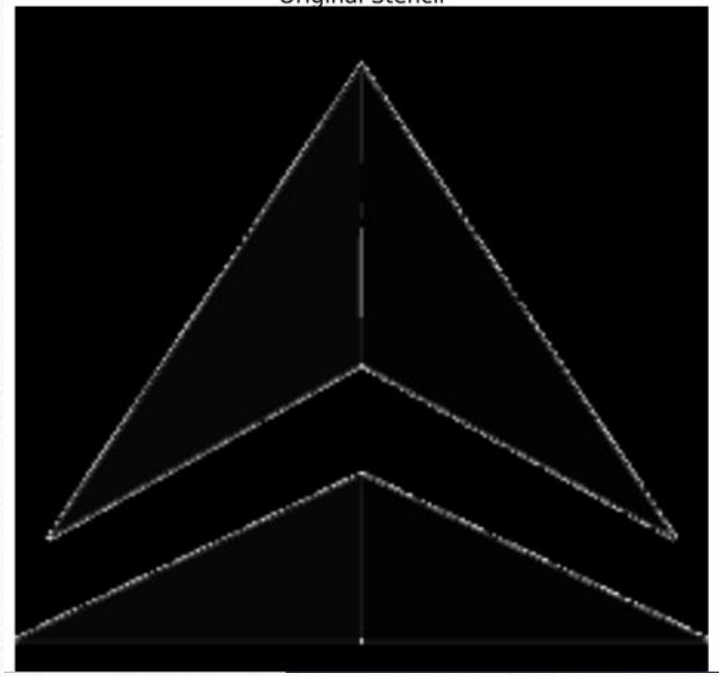
Original Stencil

Twitter Generated Cloud



Delta Airways

Original Stencil



Twitter Generated Cloud

Data Structure

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14640 entries, 0 to 14639
Data columns (total 15 columns):
tweet_id                      14640 non-null int64
airline_sentiment             14640 non-null object
airline_sentiment_confidence  14640 non-null float64
negativereason                9178 non-null object
negativereason_confidence     10522 non-null float64
airline                       14640 non-null object
airline_sentiment_gold        40 non-null object
name                          14640 non-null object
negativereason_gold           32 non-null object
retweet_count                 14640 non-null int64
text                          14640 non-null object
tweet_coord                   1019 non-null object
tweet_created                 14640 non-null object
tweet_location                9907 non-null object
user_timezone                 9820 non-null object
dtypes: float64(2), int64(2), object(11)
memory usage: 1.7+ MB
```

Findings

Random Forest

```
[[1715  109   46]
 [ 328  239   47]
 [ 134   64  246]]
              precision    recall  f1-score   support

    negative       0.79      0.92      0.85      1870
     neutral       0.58      0.39      0.47       614
    positive       0.73      0.55      0.63       444

   micro avg       0.75      0.75      0.75      2928
   macro avg       0.70      0.62      0.65      2928
weighted avg       0.73      0.75      0.73      2928

0.7513661202185792
```

Conclusion & Recommendation

- **T**hus, we can conclude that the negative sentiment about the particular industry is too high. Thus I would like to recommend the CEO (New comer) to make prior arrangement to tackle these unfavorable situation before entering into this market. And, in case the CEO is an existing player then he should try improve his grievance redressal system, so that he can take advantage existing situation and make profit in long run through taking the advantage of uniqueness

ANY QUESTIONS