MACHINE LEARNING - WEEK-4

LEARNING OBJECTIVE OF THIS MODULE

- Supervised Learning : KNN & Naïve Bayes

- Ensemble Techniques: Bagging, Boosting, Cross-validation and SMOTE

- Text Mining & Sentiment Analysis

LEARNING OBJECTIVES OF THIS SESSION

- Text Mining

- Sentiment Analysis

- Classification

BROAD OVERVIEW

Sentiment	Tweets
Negative	@united is the worst. Nonrefundable First class tickets? Oh because when you select Global/FC their system auto selects economy w/upgrade.
	@united I will not be flying you again
Neutral	@VirginAmerica my drivers license is expired by a little over a month. Can I fly Friday morning using my expired license?
	@VirginAmerica any plans to start flying direct from DAL to LAS?
Positive	@VirginAmerica done! Thank you for the quick response, apparently faster than sitting on hold ;)
	@united I appreciate your efforts getting me home!

greatlearning
Learning for Life

Discovering people opinions, emotions and feelings about a product or service
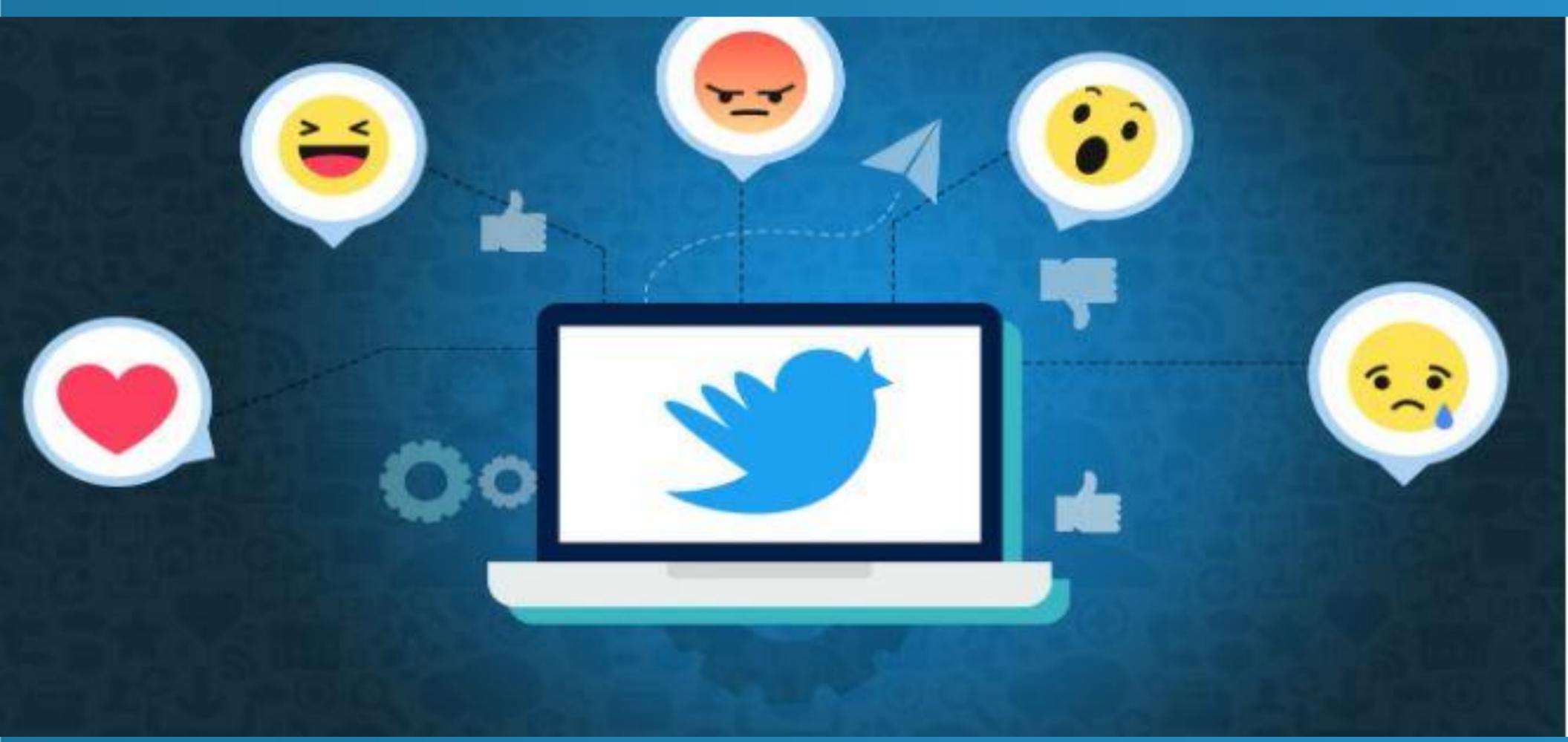
CASE STUDY- Twitter US Airline Sentiment

In recent years, Twitter has become the de facto online customer service platform. Thus, a company's image on Twitter is of central importance and this is especially true for airlines given that many tweets are travel-related in nature. In fact, research has shown that responding to tweets has revenue generating potential, drives higher satisfaction than other customer service channels, and perhaps most importantly, satisfied Twitter users spread the word.

In this case study, tweets have been gathered from Twitter to learn about people's flight experiences and give airline companies suggestions on how to make their trip more enjoyable.

The data set contains about 15,000 tweets, collected from February 2015 on Six major U.S. airline reviews. Every review is labeled as either positive, negative or neutral. First, we want to build a model to perform sentiment analysis on the data set. Secondly, more interestingly, we want to find a reason behind negative response, such as late flight, lost luggage, etc. we have been asked to classify positive, negative, and neutral tweets.

There are 14,640 rows and 15 columns. The Data include features : tweet id, sentiment, sentiment confidence score, negative reason, negative reason confidence, airline, sentiment gold, name, retweet count, tweet text, tweet coordinates, time of tweet, date of tweet, tweet location, and user time zone.

Data Science @ Work

Apply **Data Science at your workplace** to gain some instant benefits:

- Get noticed by your management with your outstanding analysis backed by data science.

- Create an impact in your organization by taking up small projects/initiatives to solve critical issues using data science.

- Network with members from the data science vertical of your organization and seek opportunities to contribute in small projects.

- Share your success stories with us and the world to position yourself as a subject matter expert in data science.

HAPPY LEARNING

Excelerate CareerPrep:
Career services by Great Learning.

G Great Learning

The CareerPrep is an organized way to learn about all of the topics asked in interviews for Data Science jobs.

- Complete Excelerate career prep and learn how to crack an interview with ease.

- Understand the different kind of roles available in the data science ecosystem along with their skill set requirements

- Build a resume that has a higher probability of getting shortlisted by recruiters

- Prepare for different data science roles by reviewing 400+ questions asked in actual interviews

- Understand and implement frameworks to solve guesstimates, logical and aptitude problems often used in interviews

- Learn from the experience of alumni who have successfully transitioned

- Speed up your path towards a transition to a Data science role.

Complete Career Prep today to speed up your path towards DS transition. You'll have a 40% better chance of getting hired than other candidates.

Note: CareerPrep is mandatory step in unlocking the job board.