ROHAN NYATI

500075940

R177219148

BATCH – 5

Assignment – 3

Q.1 Social media analytics is **the ability to gather and find meaning in data gathered from social channels to support business decisions** — and measure the performance of actions based on those decisions through social media. ... Social media analytics includes the concept of social listening.

Q.2 When we make any transaction while purchasing any product online — a good amount of people prefer credit cards. The credit limit in credit cards sometimes helps us me making purchases even if we don't have the amount at that time. but, on the other hand, these features are misused by cyber attackers.

To tackle this problem we need a system that can abort the transaction if it finds fishy.

Here, comes the need for a system that can track the pattern of all the transactions and if any pattern is abnormal then the transaction should be aborted.

Q.3 A word embedding is **a learned representation for text where words that have the same meaning have a similar representation**. It is this approach to representing words and documents that may be considered one of the key breakthroughs of deep learning on challenging natural language processing problems.

Q.4 **Neural machine translation** (NMT) is an approach to machine translation that uses an artificial neural network to predict the likelihood of a sequence of words, typically modeling entire sentences in a single integrated model. Attention mechanisms are being increasingly used to improve the **performance** of Neural Machine Translation (NMT) by selectively focusing on sub-parts of the sentence during translation.

Q.5 The k-nearest neighbors (KNN) algorithm is **a simple, supervised machine learning algorithm** that can be used to solve both classification and regression problems. It's easy to implement and understand, but has a major drawback of becoming significantly slows as the size of that data in use grows.

Q.6 **Principal Component Analysis (PCA)** is a statistical procedure that uses an orthogonal transformation that converts a set of correlated variables to a set of

uncorrelated variables. PCA is the most widely used tool in exploratory data analysis and in machine learning for predictive models. Moreover, PCA is an unsupervised statistical technique used to examine the interrelations among a set of variables. It is also known as a general factor analysis where regression determines a line of best fit.

Q.7 Well, you need to know that the distinction between Bayes theorem and Naive Bayes is that **Naive Bayes assumes conditional independence where Bayes theorem does not**. This means the relationship between all input features are independent. Maybe not a great assumption, but this is is why the algorithm is called “naive”.

Q.8 Machine learning forecasting for banking enables **more accurate reporting by automating credit risk testing for both banks and customers**. By evaluating a consumer's financial history, recent transactions, and purchasing patterns, machine learning can make accurate forecasts of future spending and income. Armed with machine learning and artificial intelligence technologies, they have **the opportunity to analyze data that originates beyond the bank office**. Financial companies collect and store more and more user data in order to revise their strategies, improve user experience, prevent fraud, and mitigate risks.

Q.9 The bullwhip effect (also known as the Forrester effect) is defined as **the demand distortion that travels upstream in the supply chain from the retailer through to the wholesaler and manufacturer due to the variance of orders** which may be larger than that of sales. The bullwhip effect is caused by **demand forecast updating, order batching, price fluctuation, and rationing and gaming**. ... Price fluctuations due to inflationary factors, quantity discounts, or sales tend to encourage customers to buy larger quantities than they require.

Q.10 At the most basic level, a chatbot is **a computer program that simulates and processes human conversation (either written or spoken)**, allowing humans to interact with digital devices as if they were communicating with a real person.

Insurers use machine learning **to predict premiums and losses for their policies**. Detecting risks early in the process enables insurers to make better use of underwriters' time and gives them a huge competitive advantage.

The crime rate is ever-growing and so is technology. I thought of combining them to produce meaningful insights. Though, there is a plethora of work on London & Chicago crime analysis but few on Indian records. So, I found a data-set from data.world for this study. The objective of this project is to find out which state has the highest crime rate and which is the crime that is committed. Without further ado, let’s get started.

Q.11 Apriori algorithm refers to the algorithm which is used to calculate the association rules between objects. It means how two or more objects are related to one another. In other words,

we can say that the apriori algorithm is an association rule leaning that analyzes that people who bought product A also bought product B.

The primary objective of the apriori algorithm is to create the association rule between different objects. The association rule describes how two or more objects are related to one another. Apriori algorithm is also called frequent pattern mining. Generally, you operate the Apriori algorithm on a database that consists of a huge number of transactions. Let's understand the apriori algorithm with the help of an example; suppose you go to Big Bazar and buy different products. It helps the customers buy their products with ease and increases the sales performance of the Big Bazar.