**TASK-01**

REAL WORLD APPLICATIOND OF DATA SCIENCE

1.HEALTHCARE

2.E-COMMERCE

3.AUTOMATION

4.SENTIMENT ANALYSIS

5.SPORTS

**HEALTHCARE:**

In Healthcare applications we can predict that the customer was affect by the disease or not with the help of Machine Learning.

**E-COMMERCE:**

In an ecommerce Application we collect the user data and build the recommendation system as what the user need.

**AUTOMATION:**

Machine Learning is used in the automation system for the decision making process . For example Tesla.

**SENTIMENT ANALYSIS:**

In socialmedia Machine Learning is used for the content recommendation as the users intrest.

**SPORTS:**

In Sports it is used for Teams analyze player performance, injury risks, and game strategies using big data .

**TASK-02:**

DEFINITIONS OF AI/ML/DL

**AI(ARTIFICIAL INTELLIGENCE):**

 AI refers to machines that can perform tasks that normally require human intelligence, such as learning, problem-solving, and reasoning.

**ML(MACHINE LEARNING):**

Machine Learning uses algorithms to develop a model for decision making .ML is the subset if the AI(Artificial Intelligence).

**DL(DEEP LEARNING):**

Deep learning is the subset of the ML(Machine Learning).DL uses Neural networks to handke the data. The neural networks are ANN(ARTIFICIAL Neural Network),RNN(Recurrent Neural Network),CNN(Convolutional neural network) etc.. DL is mostly used for image processing

**TASK 3:**

**Data Analyst**

Extracts and cleans data from various sources, performs exploratory data analysis (EDA), and generates reports and visualizations to communicate insights to stakeholders.

**Machine Learning Engineer**

Develops and trains machine learning models, collaborates with data scientists to implement models into production systems, and ensures the models are scalable and efficient.

**Data Engineer**

Designs, builds, and maintains data pipelines and infrastructure, ensures data quality and integrity, and enables efficient data access for analysis and reporting.

**TASK 4:**

**Diabetes Prediction in Healthcare:**

**1. Problem Definition:**

To predict the diabetes for the patient using machine learning

**2. Data Collection:**

Collect the patient data from any resource that include age,Blood pressure,glucose level ,gender,BMI and ect..

**3. Data Cleaning:**

Remove the duplicates and irrelevant coloumns.Handle missing values and replace it bye their mean value.also need to do standardization .by the use of encoding technique converts the categorical values into continuous values.

**4. Exploratory Data Analysis (EDA):**

Visualize data using histograms, scatter plots, and box plots. Identify correlations between features and diabetes occurrences. Identify correlations between features and diabetes occurrences.

**5. Feature Engineering:**

Create new features based on domain knowledge (e.g., age group, BMI category). Encode categorical variables for example one-hot encoding for gender. Normalize numerical features to ensure consistent scaling.

**6. Model Building:**

Select the Algorithm such as logistic regression,random forest and so on.split the data set into training set and testing set .Training data set should be 70% and testing should be 80%.

Next do the Hyperparameter Tuning:Optimize model parameters using techniques like Grid Search or Random Search.

**7. Model Evaluation:**

Evaluate the model with metrices like precision,recall,F1 score and accuracy score. Avoid overfitting in the model and if the accuracy score is less than 70% then neeed to retrain the model.

**8. Deployment:**

Deploy the model in the product environment which is able to use by the user.