MACHINE LEARNING

ASSIGNMENT-1

**REPORT**

**REGISTRATION NUMBER:** BL.EN.U4CSE21119

**NAME:** MAHI KOLLI

1) **Discuss the importance of rank of an observation matrix in model building for classification .**

The rank of an observation matrix is essential in model building for classification for the following reasons:  
  
Dimension Reduction: Sufficient information to discriminate between classes is reflected by a high-rank matrix. Suggesting that there are enough linearly independent features to capture...  
  
Avoiding Overfitting: Especially when the number of samples is low, high-feature matrices can cause overfitting in classification models. By choosing relevant features or carrying out dimensionality reduction (including PCA), you may lessen the rank and address this problem.  
  
Feature Selection: Classification's informative features can be identified with the aid of the rank. Features that don't offer useful information for separating classes are implied by low-rank matrices.  
  
Multicollinearity: High-rank matrices with collinear features (multicollinearity) may result in classification model instability. Essential is selecting uncorrelated features to reduce multicollinearity.

2)**Discuss on regression (Ex: A2) and classification (Ex: A3) tasks. How would you differentiate between them.**

Discuss regression (Ex: Tasks like A2) and classification (Ex: A3) are part of. Differentiating between them, how do you?  
  
Regression: Continuously quantifiable values are forecasted through the regression supervised learning task. Predicting product costs, A2 presents a regression challenge. Finding relationships between input features and continuous targets is the main goal of regression models.  
  
Classification: Discrete class labels or categories, classification tasks involve predicting them through supervised learning. Classifying customers based on their purchasing habits, A3 aims to distinguish between RICH and POOR. Predefined categories or classes are what classification models assign input data points to.  
  
Key Differences:  
  
Output: While in regression, the output is a continuous variable, compared to classification, where the output is a categorical variable.  
Goal: Establishing the link between variables, regression is concerned with making numerical forecasts.  
Evaluation: Classifications models are assessed using metrics including accuracy, precision, recall, F1-score, and confusion matrix.

3)**Observing the stock data provided, record your suggestionsto build a system that may be able to predict the price and Change % into future.**

How might we design a system capable of forecasting future price changes?  
  
To build a system for predicting stock prices and change percentages into the future, consider the following suggestions:  
  
Time Series Analysis: Time series analysis techniques enable modeling of historical stock prices. Trends, seasonality, and autocorrelation patterns can be identified when dealing with this.  
  
Feature Engineering: Features such as moving averages and technical indicators, including RSI and MACD, can be developed from news articles or social media.  
  
Model Selection: From autoregressive models (ARIMA) to machine learning algorithms (decision trees, random forests, or gradient boosting) and deep learning models (LSTM or GRU), experiment with diverse predictive models.  
  
Cross-Validation: Cross-validation techniques help evaluate models without succumbing to overfitting.  
  
External Data: External data sources affecting stock prices include economic indicators, company news, and global events.  
  
Evaluation: Assessing a model's effectiveness relies heavily on metrics like Mean Absolute Error and testing it against historical data.  
  
Ensemble Methods: To improve prediction accuracy and resilience, ensemble methods involve blending forecasts from various models.  
  
Predicting Change %: Through regression modeling, predict future stock price changes by estimating their values.  
  
Risk Management: In case of incorrect predictions, risk management strategies should be employed to mitigate potential losses.  
  
Continuous Learning: Update the model with fresh information to accommodate evolving market circumstances.  
  
Regulatory Considerations: Building and deploying financial prediction models, regulatory constraints and ethical considerations must be taken into account.  
  
A system for predicting stock prices calls for proficiency in finance and data science.