Academic Year	Module	Report Number	Report Type
2025	AI and ML	03	Answer to the provided question

(AI and ML)

Question and Answer

Student Name: Anup Dangi

Student id: 2333319

Group: L6CG7

Module leader: Simon Giri

Tutor: Durga Pokharel

Submitted on:

Table of Contents

1. You are a Data Scientist at Esewa, Nepal's leading digital payment platform.

1.1. Customer Segmentation:

Esewa has significance for boosting user engagement and personalizing product offerings because of its millions of users, each of whom has unique purchasing and activity patterns. Esewa uses clustering techniques like K-Means or DBSCAN to identify different client categories according to transactional frequency, usage trends, and service selection. Through this study, Esewa was able to pinpoint distinct consumer profiles that would enable tailored marketing campaigns, tailored loyalty programs, and focused financial services like insurance and microloans. Lastly, it can boost brand loyalty and consumer happiness, which can support growth even in the face of competitive pressures.

1.2. Fraud pattern detection:

To safeguard the integrity of the financial environment at Esewa, it is crucial to spot unusual or fraudulent activity in transactions that deviate from the normalcy of user profiles. Auto-encoders can assist in identifying suspicious transactions that indicate fraudulent activity by using anomaly detection techniques, such as Isolation Forest. These techniques are made possible by real-time analysis of the transaction data, which enables them to identify suspicious transactions further, send out timely alerts, and take prompt action, such stopping or evaluating a transaction, protecting users and minimizing any financial losses. A more secure and reliable transaction platform results from such actions, which also improve security and foster user trust.

2. Short Question:

2.1. Overfitting:

When a model becomes too complicated, it may become difficult to remember the specifics of the training dataset, like noise and irregularities, and it may not be able to explain strong trends, which could result in poor generalization to new data. However, underfitting is linked to such a model, which performs badly on training and test data samples because it is too simplistic to identify patterns in the data.

Example of overfitting: A deep neural network can achieve 99% accuracy in training on a tiny sentiment analysis dataset, but its performance drops to 60% on the test set. This suggests that the model performs poorly with unknown data because it has over-learned the training set, which includes errors and unpredictable fluctuations.

Example of underfitting: The goal of a linear regression model is to explain non-linear relationships in customer purchasing. Reduced accuracy on the train and test data results from the model's inability to represent non-linear connections with a straight line.

A model's effectiveness sharply declines when it either overfits or underfits the data. An underfitting model does a poor job of using the available data to generate useful predictions, while an overfitting model does well during the training phase but will function poorly in the actual world. Achieving the ideal balance in model complexity is essential for both optimal performance and well-fitting data.

2.2. Neural Network Architecture:

2.2.1. Differences between CNN and RNN:

Aspect	**CNN(Convolutional Neural Network)**	**RNN(Recurrent Neural Network)**
Date type	Spatial data (e.g., images)	Sequential data (e.g., text, time-series)
Architecture	Uses convolutional layers with filters to extract features.	Uses recurrent connections to maintain temporal context
Information Flow	Processes data in parallel	Processes data in sequence (one step at a time)
Training Challenges	High computation on large images	Vanishing/exploding gradient issues
Solutions to Challenges	Pooling, batch normalization, data augmentation	Use of LSTM, GRU, gradient clipping
Key Advantage	Efficient at detecting spatial hierarchies	Captures temporal patterns and dependencies

2.2.2. Training Challenges:

- Vanishing gradient: The vanishing gradient problem arises when gradients drastically decrease during backpropagation, making it challenging to efficiently update weights in deep neural networks. The exponential decay of gradients during backpropagation across the network makes the problem more severe in the case of conventional feedforward neural networks, which have numerous layers. This problem is resolved by sophisticated network designs like LSTM networks or GRUs, which are special-purpose algorithms that maintain gradients over lengthy durations. Adopting ReLU, which does not saturate for input greater than zero, is advantageous since it preserves gradient flow, enabling quick and efficient training of deep networks.
- Overfitting: A model experiences overfitting, which causes it to perform poorly on new data, if it learns the statistical abnormalities or peculiarities of the training set using the actual patterns. Using dropout, which involves randomly turning off some neurons during training to increase the model's generalizability, is one technique to prevent overfitting. Another popular technique is early stopping, which prevents the model from learning noise by halting training as soon as performance on the validation set starts to deteriorate. Additionally, by adding new viewpoints to the training set, data augmentation broadens the dataset's diversity and enables the model to perform well on unseen data.