UNIVERSITY PARTNER

UNIVERSITY OF WOLVERHAMPTON

HERALD COLLEGE KATHMANDU

Artificial Intelligence and Machine Learning

(6CS012)

Question And Answers

Student Id : 2227486

Student Name : Nayan Raj Khanal

Group : L5CG4

Tutor : Ms. Sunita Parajuli

Lecturer : Mr. Siman Giri

Date : 12th May, 2024

Table of Contents

1. ML Applications:

Question:

As a ML engineer, working at fin-tech company(e.g: khalti or esewa) you are planning to implement machine learning solutions, list out some of the challenges you might face and how would you address these challenges.

Solution:

For the financial industry challenges can be:

- Scalability requires vast data and when gathering and managing the financial data, data sensitivity is paramount.
- Being highly regulated, it must manage data while complying to strict regulations and providing security measures to protect sensitive information.
- Lack of domain-specific knowledge holds back ground-breaking solutions which is further worsened by the fact of the shortage of AI talent.
- Increase of AI use increases the risks of cybersecurity.
- As AI models are complex it can be difficult to build customers trust as it becomes harder to explain and without explainability there's risk of losing customers trust.

Solutions to these hurdles can be:

- Since many don't have knowledge about the utilization and limitations of AI, industry should invest in education their employees.
- Invest in hardware, software and personnel who can navigate the system effectively and support the implementation of AI,
- Implement the best security protocols as this industry is highly regulated and customers trust is based upon it.
- AI systems should remain accountable for the decisions and must be made to explain their steps when necessary keeping trust and ethics paramount.

(Fusemachines, 2023)

2. Training of Machine Learning:

Question:

Explain the concept and use of Cost/Loss/Error function in training of machine learning. Briefly explain the error function used to train Linear Regression.

Solution:

In Machine Learning, loss function is a measure that tells us how well the machine learning algorithm predictions is matching the actual target values. It is a difference in the form of a single real number which measures the performance of a model for any given data. (Crypto1, 2024)

The role of loss function includes:

- Providing difference between actual and predicted value offering a clear metric to evaluate a model.
- Directing the model to adjust parameters iteratively guiding it to reduce loss and improve predictions.
- Generalizing model to new data by balancing bias and variance.

In linear regression, the loss function commonly used is the Mean Squared Error (MSE) or L2 Loss. MSE calculates the average squared difference between the predicted values and the actual values.

The formula for MSE is:

$MSE = (1/n) * \Sigma(y_i - \bar{y})^2$

Where:

- n is the number of samples in the dataset
- y_i is the predicted value for the i-th sample
- $\bar{y}$ is the target value for the i-th sample

MSE is used particularly in scenarios where outliers need significant penalization, only when outliners are due to meaningful signals and not noise.

(Alake, 2023)

3. Overfitting:

Question:

What are overfitting and underfitting? Why is it a problem in machine learning? Explain the concepts with the help of examples.

Solution:

A student mugs up every word in a textbook without understanding the concepts. In exam time, if the questions are slightly twisted the student struggles. This is overfitting, model tries to learn every single data point essentially mirroring the training data. The model memorizes the training data, leading to high accuracy in training set but low in unseen data. The overfitted model has low bias and high variance.

On the other hand, a student only reads the summary of each chapter, missing out on the details, they might have general idea but fail when faced with specific question. This is underfitting, model is too simplistic to capture the underlying patterns of the data. Here, the model is not able to learn enough from the training data, resulting in poor performance on both the training and new data. An underfitted model has high bias and low variance.

Both Overfitting and underfitting are bad as they lead to poor generalization. We aim to build models that can accurately predict on unseen data. Overfitting and underfitting ruin this aim by either capturing too much noise or too little information from the training data, respectively.

(ScribbleData, 2023) (Saxena, 2023) (Jaiswal, 2024)

4. Neural Network Architecture

Question:

Explain the difference between a convolutional neural network (CNN) and a recurrent neural network (RNN). When would you use one over the other? Explain in brief, the challenges faced while training the deep learning models and how can they be addressed?

Solution:

Key differences between CNN and RNN are:

Aspects	Convolutional Neural Network (CNN)	Recurrent Neural Network (RNN)
Task	Image processing tasks	Sequential data tasks
Input Data	Fixed size (e.g, images)	Variable-length (e.g, text, speech)
Structure of data	Hierarchical features	Sequential influence
Handling of data	Convolutional feature extraction	Recurrent memory maintenance
Variable-length	Not designed for variable length	Handles variable length
Training Complexity	Less complex for images	Complex due to vanishing gradients
Use Case	Grid-like data tasks	Order and context-dependent tasks

Table 1. CNN VS RNN (Shahrour, 2023)

Challenges faced while training deep learning models:

Overfitting: Models may perform well on training data but fail to generalize to unseen data.

Vanishing or Exploding Gradients: During backpropagation gradients may become too small (vanishing) or too large (exploding), making training difficult.

Slow Training and Computational Resources: Deep learning models often require large amounts of data and computational resources for training.

Hyperparameter Tuning: Choosing the right set of hyperparameters (learning rate, batch size, etc.) can significantly impact model performance.

(Found, 2023)

Addressing these challenges:

Regularization Techniques: Methods like dropout, L1/L2 regularization help prevent overfitting.

Gradient Clipping: Limit the size of gradients to prevent gradients from vanishing or exploding.

Model Architectures: Using proven architectures as starting point, like CNNs for image classification or LSTMs for sequence data.

Transfer Learning: Utilize pre-trained models and fine-tuning on specific tasks to reduce training time and resource requirements.

Grid Search: Search for best hyperparameters by trying out all possible combinations.

(Found, 2023)

5. References

Alake, R., 2023. *Loss Functions in Machine Learning Explained*. [Online]
Available at: https://www.datacamp.com/tutorial/loss-function-in-machine-learning
[Accessed 6 May 2024].

Crypto1, 2024. *Gradient Descent Algorithm: How does it Work in Machine Learning?*. [Online]
Available at: https://www.analyticsvidhya.com/blog/2020/10/how-does-the-gradient-descent-algorithm-work-in-machine-learning/#:~:text=Gradient%20descent%20is%20an%20optimization%20algorithm%20used%20to%20minimize%20the,opposite%20direction%20of%20the%20gradient.
[Accessed 6 May 2024].

Found, L. a., 2023. *Common Problems When Training Deep Learning Models and How to overcome Them*. [Online]
Available at: https://medium.com/@lostandfound2654/common-problems-when-training-deep-learning-models-and-how-to-overcome-them-e37d0ac0a13b
[Accessed 6 May 2024].

Fusemachines, 2023. *AI in Fintech: Major Challenges and Solutions in 2023*. [Online]
Available at: https://insights.fusemachines.com/ai-in-fintech-major-challenges-and-solutions-in-2023/#h-major-ai-in-fintech-challenges
[Accessed 6 May 2024].

Jaiswal, S., 2024. *Overfitting and Underfitting in Machine Learning*. [Online]
Available at: https://www.javatpoint.com/overfitting-and-underfitting-in-machine-learning
[Accessed 6 May 2024].

Saxena, S., 2023. *Underfitting and Overfitting in Machine Learning*. [Online]
Available at: https://www.analyticsvidhya.com/blog/2020/02/underfitting-overfitting-best-fitting-machine-learning/
[Accessed 6 May 2024].

ScribbleData, 2023. *Overfitting and Underfitting in ML: Introduction, Techniques, and Future*. [Online]
Available at: https://www.scribbledata.io/blog/overfitting-and-underfitting-in-ml-introduction-techniques-and-future/
[Accessed 6 May 2024].

Shahrour, A., 2023. *CNN, RNN, and LSTM*. [Online]
Available at: https://medium.com/@abdalrahman_shahrour/what-is-the-difference-between-cnn-rnn-and-lstm-ee24c7f3b3fd
[Accessed 6 May 2024].