**PART I: AI Deep Learning (20 Points)**

**Question 1.1:**

**--) Provide an overview (at a minimum of 2 pages, including images) of the history of artificial intelligence, including its sub-fields, machine learning and deep learning.**

**Answer:**

Artificial intelligence is a vast field which is rapidly changing today’s world. It consists of two major sub field which are:

* Machine Learning
* Deep Learning

The below diagram shows the same

A blue rectangular box with white text and blue rectangles with white text

Description automatically generated

**Artificial Intelligence:**

* Any technique that enables computers to mimic human behaviour using logic, if then rules, decision trees and machine learning including deep learning.

**Machine learning:**

* Ability to learn without explicitly being programmed.
* A subset of Artificial Intelligence which includes abstruse statistical techniques that enable machine to improve at tasks with experience. This category includes deep learning.

**Deep learning:**

* Learn underlying features in data using neural networks
* The subset of machine learning composed of algorithms that permit software to train itself to perform tasks, like speech and image recognition by exposing multilayered neural networks to vast amount of data.

**History of Artificial Intelligence:**

The below figure shows the timeline of the evolution of Artificial Intelligence, Machine Learning and Deep Learning.

A diagram of machine learning

Description automatically generated

Fig: AI evolution timeline (Source: Week 2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach)

* **Universal Turing Machine(1930s):** The Universal Turing Machine was conceptualized by British logician Allen Mathison Turing in 1930s. During World War II, Allen Turing with his team could successfully decode the Enigma which was used by the Genrmans to send communications to their Army. This was carried out on the concepts of Universal Turing Machine.
* **Perceptron by Frank Rosenblatt(1958):** Frank Rosenblatt introduced the Perceptron. A single layer neural network which can learn and make simple decisions.
* **ELIZA by Joseph Weizenbaum (1965):** ELIZA was an early example of natural language processing program. It used pattern matching techniques to involve in conversation. It depicted human like interaction with computer.
* **General Problem Solver (GPS) by Allen Newell and Herbert A. Simon (1969):** GPS demonstrated a problem-solving approach in AI using heuristics and search algorithms.
* **Neural Networks Resurgence(1987):** The backpropagation algorithm lead to multilayer perceptrons and neural networks.
* **Reinforcement Learning Breakthrough (1992):** Reinforcement learning gets popular as an effective AI technique, algorithms like Q-learning and TD-Gamon, used to advancements in autonomous decision-making.
* **Deep Blue vs. Garry Kasparov Chess Match (1996):** IBM's Deep Blue supercomputer plays against chess grandmaster Garry Kasparov the world for the first time witnesses the potential of AI in strategic games and machine intelligence.
* **GPT-2 Language Model Release (August 2019):** OpenAI releases GPT-2, an advanced text generation model. For the first time, the world raises concern about the misuse of AI generated content.
* **Waymo's Fully Autonomous Taxi Service Launches (October 2020):** A breakthrough of using AI for self-driving car.
* **ChatGPT Debut, November 2022:** ChatGpt takes the world by a storm and suddenly everybody is talking about AI and riding the AI wave.

**Question 1.2:**

**--) Provide an overview (at a minimum of 1.5 pages, including images) of deep learning, including (but not limited to) the relationship between deep learning and machine learning, artificial intelligence.**

**Answer:**

**Deep Leaning Overview:**

Deep learning is a machine learning technique that uses multilayer neurons which is called as Artificial Neural Network. As the Artificial Neural Network are trained in multiple layers it is called “deep learning”. The building block of deep learning is a perceptron which is inspired by human brain cell or neuron. Deep learning models are designed to work like human brain and can efficiently recognize patterns, process information and take decisions.

A diagram of a nerve cell

Description automatically generated

Figure: Human neuron (Source: Week 2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach)

A diagram of a diagram

Description automatically generated

Figure: McCulloch–Pitts neuron (Source:Week2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach)

A diagram of a diagram

Description automatically generated

Figure: Perceptron (Source: Week2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach)

Some crucial components of Deep learning are:

* **Neural Networks:** it consists of artificial neurons which receive the input process a linear transformation and pass the result through a non-linear activation function.

Neural Network is arranged in layers and each layer is connected to the other layers.

* + Input Layer: takes the input
  + Hidden Layer: intermediate layer between input and output. Transforms the input data to be used by the output layer.
  + Output layer: provides the final output which may be prediction or classification.
* **Activation Function:** Common Activation function used are ReLU, sigmoid and tanh.
* **Training the Deep Learning Models:** 
  + **Foreword propagation:** input data is passed is passed through the network layers which generates the output.
  + **Loss function:** Measures the performance of the output by measuring the difference between the predicted and actual output.
  + **Backward Propagation:** Once the loss function is calculated, backward propagation comes into action to minimize the loss by adjusting the weights of the network.

**Relationship between deep learning and machine learning, artificial intelligence:**

**Artificial intelligence** is a broader field consisting of techniques that enables the computers to copy human intelligence. Rule based systems like Machine Learning, Natural Language Processing and Robotics are considered a part of Artificial Intelligence.

**Machine Learning:** It is a subset of Artificial Intelligence which used algorithms like decision trees, support vector machines and clustering to learn from data and make decisions. Machine Learning algorithms can train on comparatively smaller data and the results may not have a very high accuracy. It can make simple linear correlations. Can train on a CPU which means it needs comparatively lower processing power.

**Deep Learning:** A very powerful subset of Machine Learning which runs on multilayer neural networks and trans on comparatively large datasets and provides highly accurate prediction or classification results. It can make non-linear complex correlations. Needs a GPU which means needs enormous processing power.

A diagram of machine learning

Description automatically generated

Figure: Relationship between Artificial Intelligence, Machine Learning and Deep Learning

**Question 1.3:**

**--) Explain (at a minimum of 1.5 pages, including images) why Deep Learning is very popular in recent years.**

**Answer:**

Deep learning has gained immense popularity in the recent years the below mentioned points describe it in detail:

* **Availability of Large Datasets and Big Data ecosystem:** In today's world we have enormous amount of data available starting from telemetry logs to images and videos to large parquet files. It is not only about the data but also the capability to process it efficiently with big data and cloud ecosystems. The definition of large data has changed. Few decades ago, a few GB file would be a large dataset and it would be difficult to process or handle, but now even files in TB size can be processed efficiently.
* **Computational Power:** Deep neural network requires substantial computational power. With the advancement in hardware technology, now we have access to GPU (Graphics Processing Unit) and TPU (Tensor processing Unit) along with cloud computing. GPU and TPU can perform parallel computing in a highly efficient manner which is required by neural network. Availability of such infrastructure has made the deep neural network popular and accessible lately.
* **Powerful framework and libraries:** Some powerful and user-friendly frameworks such as TensorFlow and Keras has made it easily accessible to the practitioners and researchers. Community and ecosystem around these systems have made it possible for fast paced information sharing and collaboration. The overall growth in the tech ecosystem like cloud VMs have played the role of gamechanger for Deep Neural Networks.

The below images show the application of Deep Neural Network

A couple of people riding a scooter

Description automatically generated

Figure: Application of image recognition using Deep neural network

Source: Week 2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach

**A music notes and sound waves

Description automatically generated**

Figure: Use of Deep Neural Network for generating music.

Source: Week 2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach

A diagram of a variety of types of plants

Description automatically generated with medium confidence

Figure: Application of Deep Neural Network in various fields

Source: Source: Week 2\_01\_ai\_dl\_deep\_learning\_modern\_AI\_approach

With these various applications and high level of accuracy, it is natural that Deep Learning has gained popularity in the recent years.

**PART II: MLPs (Fully Connected Neural Networks) with Keras (50 Points)**

**TO-DO**

**Build, train, and evaluate a deep neural network MLP that has two layers. The training is done on the dataset pima\_diabetes.csv using the Keras sequential model in a Jupyter Notebook document.**

**Answer:**

The MLP is designed with the below listed specifications:

**Input Layer:** 8 neurons

**Hidden Layer 1:** 12 neurons

**Hidden Layer 2:** 8 neurons

**Output Layer:** 1

**Training Accuracy:** 0. 8510242104530334

**Evaluation Accuracy:** 0.7134825646877289

The below diagram shows the MLP used in the code.

A diagram of a network

Description automatically generated

Figure: MLP used in the code

The below screenshot displays the training and evaluation accuracy from the code output.

A screenshot of a computer

Description automatically generated

**Comparing the Training Accuracy and Evaluation Accuracy:**

**Training Accuracy: 0.8510242104530334**

**Evaluation Accuracy: 0.7134825646877289**

The code output shows that the Training Accuracy is higher than the Evaluation Accuracy.

A screenshot of a computer

Description automatically generated

This means that can provide an accuracy of about 71.34% on a new dataset. The difference between the training and evaluation accuracy is also known as **generalization gap.** There may be many reasons for this.

**Explaining the gap between Training Accuracy and Evaluation Accuracy:**

The generalization gap can be explained by the below listed points:

* **Overfitting:**

Overfitting is when the model predicts too well, or the accuracy is very high. This happens when the model learns the training data too well.

Overfitting Indicators:

* High Accuracy on training data
* Drop of accuracy for validation data

**Solution:**

* Increase the size of the training set so the model can generalize better.
* Add dropout layers to prevent overfitting.
* **Data Quality:**

The quality of data plays a significant role for the model’s accuracy. A clean dataset will higher accuracy in training as well as validation.

* **Imbalanced dataset:**

A dataset is imbalanced when there is more data for one class than the other. This may lead the model to give high accuracy for the majority class and perform poorly for the minority class. This may lead to a high training accuracy and a drop in validation accuracy.

**Solution:**

* + Resampling for the minority class.
* Adjust class weights for training to handle imbalance.
* **Insufficient Data:**
* The dataset is not large enough to capture any patterns.
* Small dataset tends to overfit.
* The dataset also does not have many features to capture the details.

**Compare the accuracy level from the evaluation process, i.e., accuracy\_evaluation, obtained in this project (MLP on pima\_diabetes.csv) with the accuracy level from the evaluation process of the project discussed in the lecture (MLP on Iris.csv)**

**Answer:**

The results for Iris.csv are shown below:

**Training Accuracy: 0.9500**

**Evaluation Accuracy: 0.9600000023841858**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

The results for pima\_diabetes.csv are shown below:

**Training Accuracy: 0.8510242104530334**

**Evaluation Accuracy: 0.7134825646877289**

**A screenshot of a computer

Description automatically generated**

We can see that Iris.csv has a higher training as well as evaluation accuracy than the pima\_diabetes.csv.

The reasons for this difference can be:

* **Data Set size:** 
  + Iris.csv is a very small dataset with only 150 records and only six features whereas the pima\_diabetes.csv has more data with 768 rows and 9 features.
* **Overfitting:**
  + Smaller dataset with limited features tends to overfit as there not much scope to explore the underlying trends as the data is extremely limited. Which we can see for the Iris.csv.
  + Pima\_diabetes.csv has comparatively more features and records which lowers the accuracy of the model both for training and validation.

**Note: The jupyter notebook for pima\_diabetes.csv and Iris.csv are submitted as part of this assignment and can be referred for more information.**

**PART III: Redesign the MLP (30 Points)**

**TO-DO**

**To improve the performance of the MLP on the dataset pima\_diabetes.csv, it is assumed that the student plans to use the trial and error approach experimenting with a new design of the MLP. There are many ways to redesign a neural network.**

**--) First, based on the knowledge of the deep neural network MLP and using critical thinking, the student redesigns the MLP neural network, then build, train, and evaluate the redesigned MLP to find out if it produces a higher accuracy level.**

**--) Using MS PowerPoint or Draw Tool in MS Word to draw the diagram of the redesigned neural networks with all the layers, the neurons, and the feed-forwarding connections.**

**--) Redo all the steps of the project “MLP on pima\_diabetes.csv” in another Jupyter Notebook document.**

**Answer:**

I have redesigned the MLP neural network with the below details:

**Input Layer:** 8 neurons

**Hidden Layer 1:** 12 neurons

**Hidden Layer 2:** 8 neurons

**Hidden Layer 3:** 8 neurons

**Output Layer:** 1

**Early stopping enabled**

**Training Accuracy:** 0. 8510242104530334

**Evaluation Accuracy:** 0.7134825646877289

The below diagram shows the redesigned MLP used in this part.

A diagram of a network

Description automatically generated

Figure: Redesigned MLP

The below screenshot displays the training and evaluation accuracy from the code output.

A screenshot of a computer

Description automatically generated

**Redesign changes:**

**Additional Hidden Layer:** In the redesigned MLP neural network, we have added another hidden layer with eight neurons.

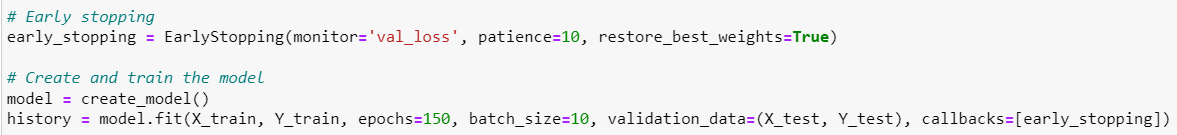
The below screenshot shows the additional third hidden layer.

A computer code with text

Description automatically generated

**Ealy stopping:** We have also added early stopping in the new design.

The below screenshot shows the early stopping.



**Performance Improvement with the Redesign in comparison with Part II:**

The redesigned model has additional hidden layer with eight neurons which increases the computation time and resource utilization but it also has a better evaluation accuracy.

We can compare the Evaluation accuracy computed earlier for the model in Part II and the redesigned accuracy.

Evaluation Accuracy for part II: 0.7134825646877289

Evaluation Accuracy for redesigned model part III: 0.7538790047168732

We can see that the redesigned model has better evaluation accuracy which means that it will predict the result at an accuracy of 75.38% for a new data. Earlier model have an accuracy of 71.34% with new data.

The redesign of MLP has helped to improve the accuracy of the model.

* **Additional Hidden Layer:** Having additional layer enhances the model’s ability to learn complex patterns from the data.
* **Ealy stopping:** When the validation performance does not improve for a certain number of consecutive epochs, early stopping stops the training. This prevents the model from overfitting.

**Note**: As part of the submission requirement, I have added three jupyter notebooks:

* Assignment3MLPs\_on\_Iris\_with\_Keras\_part\_III
* Midterm\_Assesment\_Part\_II
* Midterm\_Assignment\_Part\_III

References:

<https://www.theainavigator.com/ai-timeline>

[Deep Learning vs. Machine Learning: A Beginner’s Guide | Coursera](https://www.coursera.org/articles/ai-vs-deep-learning-vs-machine-learning-beginners-guide)