Artificial Intelligence and Machine Learning Fundamentals

CURRICULUM:

Unit I: Intelligent Agent And Uninformed Search

Introduction - Foundations of AI - History of AI - The state of the art - Risks and Benefits of AI - Intelligent Agents - Nature of Environment - Structure of Agent - Problem Solving Agents - Formulating Problems - Uninformed Search - Breadth First Search - Dijkstra's algorithm or uniform cost search - Depth First Search - Depth Limited Search

Unit II: Problem Solving With Search Techniques

Informed Search - Greedy Best First - A* algorithm - Adversarial Game and Search

- Game theory - Optimal decisions in game - Min Max Search algorithm

- Alpha-beta pruning - Constraint Satisfaction Problems (CSP) - Examples - Map Coloring – Job Scheduling - Backtracking Search for CSP

Unit III: Learning

Machine Learning: Definitions – Classification - Regression - approaches of machinelearning models - Types of learning - Probability - Basics - Linear Algebra –Hypothesis space and inductive bias, Evaluation. Training and test sets, crossvalidation, Concept of over fitting, under fitting, Bias and Variance - Regression: Linear Regression - Logistic Regression

Unit IV: Supervised Learning

Neural Network: Introduction, Perceptron Networks – Adaline - Back propagation networks - Decision Tree: Entropy – Information gain - Gini Impurity – classification algorithm - Rule based Classification - Naïve Bayesian classification - Support VectorMachines (SVM)

Unit V: Unsupervised Learning

Unsupervised Learning – Principle Component Analysis - Neural Network: Fixed Weight Competitive Nets - Kohonen Self-Organizing Feature Maps – Clustering: Definition - sTypes of Clustering – Hierarchical clustering algorithms – k-means algorithm

Course Duration: 45 Hours

Test Projects:

- Apply acquired knowledge and skills to design and develop a real-world application.
- Emphasis on Practical Problem Solving
- Work independently or in teams, demonstrating proficiency in the Artificial Intelligence Project building.

Real- world Problem Statements: **Students are required to complete any one problem and design a viable solution **

1. **Agricultural Raw Material Analysis** - You are tasked to analyze agricultural-raw-material-prices dataset over the years (EDA)
   - Find the high range and low range raw materials according to their prices.
   - high and low %Change materials
   - Identify the range of prices changed over the years.
   - Map a correlation between them using a heatmap

2. **E-commerce Sales Analysis** - You are tasked to perform an E-commerce Sales Analysis that includes the components below.
   - Analyze sales data from an e-commerce website.
   - Calculate key metrics like total sales, best-selling products, and customer purchase patterns.
   - Generate reports and visualizations to present insights for decision-making.

3. **Credit Card Fraud Detection** - You are tasked to perform Credit Card Fraud Detection to identify patterns that indicate if an applicant will repay their instalments which may be used for taking further actions such as denying the loan, reducing the amount of loan, lending at a higher interest rate, etc. This will make sure that the applicants capable of repaying the loan are not rejected. You will also develop a basic understanding of risk analytics and understand how data can be utilized to minimize the risk of losing money while lending to customers.

4. **Stock Market Forecast** - You are tasked to perform Stock Market Forecasting using linear regression.
   - Load historical pricing data into Pandas DataFrame.
   - Add technical indicators to use as features in our Linear Regression model.
   - Train a simple linear regression model using a 10-day exponential moving average as a predictor for the closing price.
   - Analyze the accuracy of the model, plot the results, and consider the magnitude of our errors.

5. **Heart Disease Prediction** - You are tasked to perform Heart Disease Prediction Using Logistic Regression. The World Health Organization has estimated that four out of five cardiovascular disease (CVD) deaths are due to heart attacks. This whole research intends to pinpoint the ratio of patients who have a good chance of being affected by CVD and predict the overall risk using Logistic Regression.

6. **Spotify Music Recommendation System** - You are tasked to perform Spotify Music Recommendation System.

   - Use the Spotify songs dataset available on Kaggle.
   - The recommender system will group relevant data points using K-Means clustering.
   - After developing the recommender system, you may deploy it as a standard Python application.

- Users can enter their favorite songs on Spotify, and your model will immediately display the most like their preferred songs as recommendations on the user's screen.

7. **Detecting Spam Emails** - You are tasked to perform Detecting Spam Emails Using TensorFlow. Implement and build a deep-learning model for Spam Detection. The model we will try to implement will be a Classifier, which would give binary outputs- either spam or ham. Steps involved –
   - Import dependencies; load and analyze the spam text data
   - Split the data into train and test sub-datasets; text preprocessing
   - Train our model using the three deep-learning algorithms
   - Compare results and select the best model
   - Use the final classifier to detect spam messages

8. **Employee Churn Prediction** - You are tasked to perform Employee Churn prediction in Python. Employee churn can be defined as a leak or departure of an intellectual asset from a company or organization. or in simple words, you can say, when employees leave the organization is known as churn. The following points help you to understand, employee and customer churn in a better way:
   - The business chooses the employee to hire someone while in marketing you don't get to choose your customers.
   - Employees will be the face of your company, and collectively, the employees produce everything your company does.
   - Losing a customer affects revenues and brand image. acquiring new customers is difficult and costly compared to retaining existing customers. Employee churn is also painful for companies in organizations. It requires time and effort to find and train a replacement.

   You are going to cover the following steps –
   - Exploratory Analysis
   - Data Visualization
   - Cluster Analysis

- Building Prediction Model
- Evaluating Model Performance

9. **Automatic Music Generation** – You are tasked to develop an end-to-end model for Automatic Music Generation.
   - Understand the WaveNet architecture and implement it from scratch using Keras.
   - Compare the performance of WaveNet versus Long Short-Term Memory for building an Automatic Music Generation model.
   - Collect as much training data as you can since the deep learning model generalizes well on the larger datasets.

10. **AI Chatbot using ChatGPT** – You are tasked to build your own AI chatbot using the ChatGPT API. From setting up tools to installing libraries, and finally, creating the AI chatbot from scratch, we have included all the small details for general users here.
    - You can build a ChatGPT chatbot on any platform, whether Windows, macOS, Linux, or ChromeOS.
    - To create an AI chatbot, you don't need a powerful computer with a beefy CPU or GPU.
    - The heavy lifting is done by OpenAI's API on the cloud. you will need Python, Pip, OpenAI, and Gradio libraries, an OpenAI API key, and a code editor like Notepad++.

11. **An End-to-End Data Science Project with ChatGPT** - You are tasked to use ChatGPT for project planning, data analysis, data cleaning and preprocessing, model selection, hyperparameter optimization, and creating and deploying web applications. You are required to use various prompts to create a project outline, write Python code, conduct research, and debug the application.
    Here is the list of tasks that we are going to follow in this tutorial:
    - Data cleaning and preprocessing. Handling missing values, converting categorical variables into numeric variables, scaling/normalizing the data, and handling any outliers or anomalies

in the data. In addition, you'll want to address the class imbalance issue by either oversampling the minority class (loans that are not fully paid) or under sampling the majority class (loans that are fully paid).

- Exploratory data analysis (EDA). Exploring the dataset to gain insights into the data, such as distributions of variables, correlations between variables, and identifying any patterns in the data.
- Feature engineering. Creating new features or transforming existing features to improve the predictive power of the model.
- Model selection. Try several models, such as decision trees, random forests, logistic regression, or support vector machines (SVM).
- Model training and evaluation. Training the selected model on the data and evaluating its performance using various metrics such as accuracy, precision, recall, and F1 score.
- Hyperparameter tuning. Fine-tuning the hyperparameters of the selected model to improve its performance.
- Creating a web app using Gradio. Once you have selected the best model, you can create a web app using Gradio.
- Deploying the web app on Spaces. Once you have created the web app using Gradio, you can deploy it on Spaces.
- Testing the web app. Test, it to make sure it is working as expected.

12. **Sentiment analysis AI system** – You are tasked to create a Sentiment analysis AI system using ChatGPT. Here are the key steps:
    - Describes the dataset to be used. Choose any one you prefer to work with or create your own new dataset.
    - Introducing the OpenAI API.
    - Installation of required libraries in Jupyter and starts using ChatGPT OpenAI API for Sentiment Analysis.
    - ChatGPT's powerful AI capabilities can be harnessed for comprehensive sentiment analysis, summarization, and actionable insights from customer reviews.