Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati

Department of Artificial Intelligence and Data Science

T.Y. B. Tech Syllabus 2025-26 (As per NEP 2020)

Syllabus: Double Minor w. e. f. AY: 2025-2026
SEMESTER-V

Double Minor in Artificial Intelligence and Data Science

SEM	Course Code	Courses Name	Teaching Scheme			Examination Scheme and Marks							Credits		
			TH	PR	TUT	Activity	ISE	ESE	TW	PR	OR	Total	TH	PR	TUT	Total
V	AI23281	Machine Learning: Concepts, Techniques, and Applications	3	2	-	20	20	70	20	20		150	3	1		4

Dept. Autonomy Coordinator
Mrs. R. S. Naik

Dept. Academic Coordinator
Mr. P. N. Shendage

HOD, AI&DS
Dr. C. S. Kulkarni

Dean Autonomy
Dr. C. B. Nayak

Dean Academics
Dr. S. M. Bhosle

Principal
Dr. S. B. Lande

V.P.K.B.I.E.T. BARAMATI
Affiliated to S.P.P.U. Pune
Id No. /PU/PN/ Engg/152/2008
PUNE (MS.)

Principal
Vidya Pratishthan's
Kamalnayan Bajaj Institute of
Engineering & Technology, Baramati
Vidyanagari, Baramati-413133

BUCKET OF DOUBLE MINOR DEGREE

DOUBLE MINOR DEGREE (only for students having CGPA >= 7.5)
AI23281: Machine Learning: Concepts, Techniques, and Applications

Vidya Pratishthan's
Kamalnayan Bajaj Institute of Engineering and Technology, Baramati
(Autonomous Institute)

AI23281- Machine Learning: Concepts, Techniques, and Applications

Teaching Scheme: **Theory: 3 Hours/Week** **Practical: 2 Hour/Week**	**Credits** **04**	**Examination Scheme:PR:20** **Activity:20 Marks** **ISE: 20 Marks** **ESE: 70 Marks** **Term Work: 20 Marks**

Prerequisites: Python Programming

Course Objectives:.

- To understand and analyze various machine learning models, their principles, and real-world applications
- To design and implement machine learning models that predict outcomes using linear, non-linear, and logistic regression techniques.
- To develop and evaluate classification models for solving binary and multiclass classification problems.
- To apply and evaluate clustering algorithms for effective segmentation and analysis of unlabeled data.
- To explore ensemble learning techniques for enhancing the performance of machine learning models.
- To understand the fundamentals of reinforcement learning and apply its algorithms to real-world applications.

Course Outcomes (COs): The students will be able to learn:

CO1: Explain and differentiate various machine learning models based on their characteristics and applications

CO2: Develop and implement predictive models using linear, non-linear, and logistic regression techniques.

CO3: Apply and evaluate classification techniques for solving binary and multiclass classification problems.

CO4: Implement and analyze clustering techniques for handling and segmenting unlabeled datasets.

CO5: Utilize ensemble learning methods to enhance model performance through the combination of multiple machine learning algorithms.

CO6: Apply reinforcement learning techniques to solve real-world decision-making problems.

Course Contents

Unit I Introduction to Machine Learning(06 Hour)

Introduction: Machine Learning, Definitions and Real-life applications, Comparison of Machine Learning with Traditional Programming, ML vs AI vs Data Science. **Learning Paradigms**: Learning Tasks- Descriptive and Predictive Tasks, Supervised, Unsupervised, Semi-supervised and Reinforcement Learnings.

Unit II: Regression (06 Hour)

Introduction: Regression, Need of Regression, Compare Regression and Correlation, Bias-Variance tradeoff, Overfitting and Under fitting. **Regression Techniques** : Support Vector Regression, Ridge Regression, Lasso Regression, Elastic Net Regression

Unit III: Classification (06 Hour)

Introduction: Need of Classification, Types of Classification (Binary and Multiclass), Binary-vs-Multiclass Classification, Balanced and Imbalanced Classification Problems , **Algorithms:** Logistic Regression for Multiclass , Decision Tree Classifier, Linear Support Vector Machines (SVM) – Introduction, Soft Margin SVM, **Kernel functions**: Radial Basis Kernel, Gaussian, Polynomial, Sigmoid

Unit IV: Evaluation Metrics (6 Hours)

Evaluation Metrics for Regression: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared ,Adjusted R-squared, **Classification Performance Evaluation**: Confusion Matrix, Accuracy, Precision, Recall

Unit V: Unsupervised Machine learning algorithms (6 Hours)

Clustering, Need of Clustering, Types of Clustering, Hierarchical clustering algorithm, K-Means clustering algorithm, Advantages and disadvantages of K-Means clustering algorithm, Elbow method.

Unit VI: Reinforcement Learning (6 Hours)

Reinforcement Learning, Need for Reinforcement Learning, Supervised vs Unsupervised vs Reinforcement Learning, Types of Reinforcement, Elements of Reinforcement Learning, Real time applications of Reinforcement learning.

Text Books:

1. **Goodfellow, I., Bengio, Y., & Courville, A.** (2016). Deep Learning. MIT Press. ISBN: 978-0262035613
2. **Raschka, S., & Mirjalili, V.** (2020). Python Machine Learning. Packt Publishing. ISBN: 978-1789955750
3. **Zhang, C., & Ma, Y.** (2021). Ensemble Machine Learning. Springer. ISBN: 978-3030985064

Reference Books:

1. Ian H Witten, Eibe Frank, Mark A Hall, "Data Mining, Practical Machine Learning Tools and Techniques", Elsevier, 3rd Edition
2. Jiawei Han, Micheline Kamber, and Jian Pie, "Data Mining: Concepts and Techniques", Elsevier Publishers Third Edition, ISBN: 9780123814791, 9780123814807
3. Shalev-Shwartz, Shai, and Shai Ben-David, "Understanding machine learning: From theory to algorithms", Cambridge university press, 2014
4. McKinney, "Python for Data Analysis O' Reilly media, ISBN : 978-1-449- 31979-3
5. Hastie, T., Tibshirani, R., & Friedman, J. (2017). The Elements of Statistical Learning. Springer. ISBN: 978-0387848570
6. Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press. ISBN: 978-0521865715

E-Resources:

1. **https://www.coursera.org/learn/machine-learning**
2. **https://www.edx.org/learn/machine-learning**
3. **https://www.kaggle.com/learn/intro-to-machine-learning**

List of Assignments

1. Use the diabetes data set from UCI and Pima Indians Diabetes data set for performing the following:
   a. Univariate analysis: Frequency, Mean, Median, Mode, Variance, Standard Deviation, Skewness and Kurtosis
   b. Bivariate analysis: Linear and logistic regression modeling
   Dataset link: https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database
2. Predict House Prices using Regression
   a. Apply Linear, Ridge, and Lasso regression to predict house prices.
   Dataset: Boston Housing Dataset
3. Predict the price of the Uber ride from a given pickup point to the agreed drop-off location.

Perform following tasks:

a. Pre-process the dataset.
b. Identify outliers.
c. Check the correlation.
d. Implement linear regression and ridge, Lasso regression models.
e. Evaluate the models and compare their respective scores like R2, RMSE.
Dataset link: https://www.kaggle.com/datasets/yasserh/uber-fares-dataset

4. Handwritten Digit Recognition
   a. Train a Logistic Regression model on MNIST for digit classification.
   Dataset: MNIST Handwritten Digits

5. Implementing K-Means Clustering on the Iris Dataset and Determining the Optimal Number of Clusters using the Elbow Method.
6. Use the Wine Quality Dataset to predict wine quality based on physicochemical properties. https://archive.ics.uci.edu/dataset/186/wine+quality
7. Stock Price Prediction using Regression
   a. Use Support Vector Regression (SVR) to predict stock market trends.
   Dataset: Yahoo Finance Stock Data
8. Use the Student Performance Dataset to predict student performance based on their study habits, family background, etc.
https://archive.ics.uci.edu/dataset/320/student+performance
9. Use the **Air Quality Dataset** to predict air quality based on several environmental features. https://archive.ics.uci.edu/dataset/360/air+quality

10. Segmenting Customers using K-Means
    a. Apply K-Means Clustering to divide customers into groups.
    Dataset: Mall Customers Segmentation

Syllabus: Double Minor w. e. f. AY: 2025-2026
SEMESTER-VI

Double Minor in Artificial Intelligence and Data Science

SEM	Course Code	Courses Name	Teaching Scheme			Examination Scheme and Marks							Credits			
			TH	PR	TUT	Activity	ISE	ESE	TW	PR	OR	Total	TH	PR	TUT	Total
VI	AI23291	Deep Learning: Foundations, Architectures, and Application	3	2	-	20	20	70	20	20		150	3	1		4

Dept. Autonomy Coordinator
Mrs. R. S. Naik

Dept. Academic Coordinator
Mr. P. N. Shendage

HOD, AI&DS
Dr. C. S. Kulkarni

Dean Autonomy
Dr. C. B. Nayak

Dean Academics
Dr. S. M. Bhosle

Principal
Dr. S. B. Lande

VPKBIET BARAMATI
Affiliated to S.P.P.U. Pune
Id No./PU/PN/Engg/152/2009
PUNE (MS.)

Principal
Vidya Pratishthan's
Kamalnayan Bajaj Institute of
Engineering & Technology, Baramati
Vidyanagari, Baramati-413133

BUCKET OF DOUBLE MINOR DEGREE

DOUBLE MINOR DEGREE (only for students having CGPA >= 7.5)
AI23291:Deep Learning: Foundations, Architectures, and Application

Vidya Pratishthan's
Kamalnayan Bajaj Institute of Engineering and Technology, Baramati
(Autonomous Institute)

AI23291- Deep Learning: Foundations, Architectures, and Application

Teaching Scheme: **Theory: 3 Hours/Week** **Practical: 2 Hour/Week**	**Credits** **04**	**Examination Scheme:PR:20** **Activity:20 Marks** **ISE: 20 Marks** **ESE: 70 Marks** **Term Work: 20 Marks**

Prerequisites: Python Programming

Course Objectives:.

- To build a strong foundation in neural networks, including their architectures, working principles, and training mechanisms.
- To analyze the evolution of deep learning, its fundamental concepts, and real-world applications.
- To explore hyperparameter tuning techniques and optimization strategies for improving deep learning model performance.
- To design and implement convolutional neural networks (CNNs) for computer vision applications.
- To develop expertise in Recurrent Neural Networks (RNNs) for sequence modeling tasks such as language processing and time series forecasting.
- To investigate advanced recurrent architectures, including LSTMs and Bi-LSTMs, for handling long-term dependencies in sequential data.

Course Outcomes (COs): The students will be able to learn:

CO1: Explain the fundamentals of neural networks, including perceptrons, activation functions, and training mechanisms...

CO2: Compare and apply deep learning techniques for various real-world applications, understanding their strengths and challenges.

CO3: Optimize neural network performance through hyperparameter tuning, gradient-based learning, and advanced optimization techniques.

CO4: Implement and train convolutional neural networks (CNNs) for solving image classification and object detection problems.

CO5: Develop and apply recurrent neural networks (RNNs) and their variants (LSTMs, Bi-LSTMs) for sequence-based data processing.

CO6: Solve complex sequence modeling problems using advanced RNN architectures and backpropagation through time (BPTT).

Course Contents

Unit I Fundamentals of Neural Networks (06 Hours)

Biological Neuron, The Perceptron(AND,OR,NOT,XOR), Deep Forward Network, Multilayer Feed-Forward Networks, **Training Neural Networks:** Backpropagation and Forward Propagation, **Activation Functions** :Linear, Sigmoid, Tanh, Hard Tanh, Softmax, Rectified Linear, **Loss Functions:**Loss Function Notation, Loss Functions For Regression, Loss Functions For Classification.

Unit II Introduction To Deep Learning Foundations (06 Hours)

Machine Learning And Deep Learning, History Of Deep Learning, Advantage And Challenges Of Deep Learning. Learning Representations From Data , Understanding How Deep Learning Works In Three Figures(Input, Hidden Layer, Output), Common Architectural Principles Of Deep Network,

Architecture Design, Applications Of Deep Learning

Unit III Hyperparameter Tuning of Neural Network (06 Hours)

Hyperparameters : Learning Rate, Regularization, Cost Functions, Error Back Propagation, Gradient-Based Learning, Implementing Gradient Descent, Vanishing And Exploding Gradient Descent, Optimization Algorithm(SGD)

Unit IV Convolutional Neural Network (CNN) (06 Hours)

CNN Architecture Overview, The Basic Structure Of A Convolutional Network- Padding, Strides, Typical Settings, The Relu Layer, Pooling, Fully Connected Layers, The Interleaving Between Layers, Training a Convolutional Network.

Unit V Recurrent Neural Networks (6 Hours)

Recurrent Neural Networks Overview, Need for RNNs, RNNs And Traditional Feedforward Networks, Basic RNN Architecture, Working of an RNN, Applications of RNNs.

Unit VI Advanced Recurrent Neural Networks and Sequence Learning (06 Hours)

Vanishing Gradient And Exploding Gradient Problem, RNN Differs From Feedforward Neural Networks, Backpropagation Through Time (BPTT), Types Of Recurrent Neuralnetworks, Bidirectional Rnns, Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (Bi-LSTM)

Text Books:

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. ISBN: 978-0262035613.
2. Patterson, J., & Gibson, A. (2017). Deep Learning. O'Reilly Media. ISBN: 978-1491925614.
3. Aggarwal, C. (2018). Neural Networks and Deep Learning. Springer. ISBN: 978-3319944623.
4. Buduma, N. (2017). Fundamentals of Deep Learning. SPD. ISBN: 978-9386551551.
5. Chollet, F. (2017). Deep Learning with Python. Manning Publications. ISBN: 978-1617294433.

Reference Books:

1. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. MIT Press. ISBN: 978-0262039246.
2. Weidman, S. (2019). Deep Learning from Scratch: Building with Python from First Principles. O'Reilly Media. ISBN: 978-1492047489.
3. Duval, F. (2020). Deep Learning for Beginners: Practical Guide with Python and TensorFlow. Independently Published. ISBN: 979-8618122373.
4. **Sutton, R. & Barto, A.** (2018). Reinforcement Learning: An Introduction (2nd Ed.). MIT Press. ISBN: 978-0262039246.

List of Assignments

1. Perceptron Model Implementation: Develop a simple perceptron-based neural network to classify logical operations (AND, OR, XOR).
2. Train a Multi-Layer Perceptron (MLP) for Digit Recognition. Use an MLP model to classify handwritten digits. Dataset: MNIST Dataset.
3. Visualizing Activation Functions Implement different activation functions (Sigmoid, ReLU, Tanh) and plot their outputs. Dataset: Synthetic data.
4. Exploring Deep Learning Architectures Implement a simple feedforward network for classification tasks. Dataset: Fashion MNIST.
5. Compare Performance of Shallow vs. Deep Networks Train 3-layer vs. 10-layer neural networks and compare accuracy. Dataset: CIFAR-10.
6. Deep Learning for Image Classification Build a deep fully connected network for image classification.

Dataset: Tiny ImageNet.

7. Implementing Regularization Techniques Compare L1 (Lasso) and L2 (Ridge) regularization on a neural network. Dataset: Diabetes Dataset.
8. Impact of Different Optimizers train a model using SGD, Adam, RMSProp and compare results.
9. Build a Basic CNN for Image Classification implement a CNN model for digit classification.
10. Sentiment Analysis using LSTM train a LSTM for IMDB movie review sentiment classification. Dataset: IMDB Dataset