Certificate Course

on

Artificial Intelligence and Machine Learning Techniques for Functional Materials

Artificial Intelligence and Machine Learning Techniques for Functional Materials

Course Code	Course Name	Credits
SPSTAIC01	**Programming Fundamentals**	**4**
SPSTAIC02	**Machine Learning-I**	**4**
SPSTAIC03	**Machine Learning-II**	**4**
SPSTAIC04	**Deep Learning Concepts**	**4**
SPSTAIP01	**Machine Learning Lab**	**2**
SPSTAIP02	**Deepa Learning Lab**	**2**
Total credits		**20**

Upon completion of the course, participants will be evaluated through four theory examinations and one mini-project. Securing a minimum of 50% marks is mandatory for the award of the certificate issued by the School of Polymer Science and Technology.

SPSTAIC01 Programming Fundamentals

Course Objectives

(1) To understand and apply core Python syntax and data structures (lists, dictionaries, loops, conditionals).
(2) To write reusable and modular code using functions and classes.
(3) To read and write files (CSV, TXT etc) for materials data analysis.
(4) Use NumPy and Pandas for manipulating materials datasets .
(5) Visualize materials data using Matplotlib and Seaborn to identify trends and patterns.

Module	Syllabus Description	Contact hours
Module 1	What is Data Science and why Data Science? Applications and Components of data Science, Why Python for Data Science? Basics of Algorithm/ Pseudocode , Program, Kinds of Programming Languages, Compilers, and Interpreters Introduction to Python, Types of IDE, Identifiers, Variables, Operators, Data Types, Conditions, Loops, regular expressions	**18**
Module 2	Introduction to Data Structures using Python: **Strings**: Introduction, functions, and operations on Strings, Application Programs on Strings. **List:** Introduction, functions and operations on List, Application Programs on Lists **Tuple**: Introduction, functions, and operations on Tuple **Dictionaries:** Introduction, functions and operations on Dictionaries, Application Programs on Dictionaries. **Sets**: Introduction, functions, and operations on Sets, Applications on Sets.	**18**
Module 3	Functions, Modules and Collections : Functions Defining and Invoking functions, Scope, Parameter types Recursive functions, Built in Functions such as enumeration, zip, sorted, map, filter and Applications Modules in Python, creating custom modules and calling them.	**18**

Module 4	Working with Materials Databases and Text: File I/O operations: Reading and Writing data from various formats, Regular Expressions, Identifiers, Quantifiers. Exploratory Data Analysis: Statistics and Probability using NumPy Data Manipulation and Exploratory Data Analysis with Pandas Matplotlib and Seaborn libraries , debugging and error handling techniques	**18**

References

1. *Matthes, E. (2019).* Python crash course: A hands-on, project-based introduction to programming *(2nd ed.).* No Starch Press.
2. *Zelle, J. (2023).* Python programming: An introduction to computer science *(4th ed.). Franklin, Beedle & Associates.*
3. Rao, N. P. (2023). *Python programming* (2nd ed.). Pearson Education India.

Video Link

https://www.youtube.com/watch?v=5zx9DI2iDsA

SPSTAIC02: Machine Learning-I

Course Objectives

1. To teach different machine learning paradigms and parameter estimation methods.
2. To enable the learner to apply the various algorithms of supervised learning in materials science contexts and analyse the overfitting and underfitting issues in machine learning models. Also equip them with various ensemble mechanisms.
3. To impart the knowledge on how to evaluate the different performance measures for analysing the performance of classification models.

Module No:	Syllabus Description	Contact hours
Module 1	Machine Learning Paradigms:- Supervised Learning, Unsupervised Learning, Semi Supervised Learning, Reinforcement Learning Parameter Estimation Methods:- Maximum Likelihood Estimation (MLE), Maximum Aposteriori Estimation (MAP), Bayes Estimation, Least Square Estimation (LS), Minimum Mean Square Estimation (MMSE), The Gradient Descent Method and its variants (Simple Gradient Descent Method, Stochastic Gradient Method, Batch Gradient Method), Illustrative examples of parameter estimation using different methods.	**18**
Module 2	Supervised Learning Models:- Regression - Regression Models (Simple Regression, Multivariable linear Regression, Polynomial Regression, Logistic Regression), Illustrative examples of designing regression models; Decision Trees - Working Principle, Decision Tree formation (ID3 Algorithm), The Classification and Regression Trees Algorithm (CART), Illustrative examples on designing decision trees; Overfitting and Underfitting :- Bias and Variance, Bias-Variance Trade-off, Regularization (Ridge, Lasso and Elastic Net Regularization techniques), Error metrics-MAE, MSE, R^2	**18**
	More Supervised Learning Methods:- Support Vector machines (SVM) - Maximum Margin Classification, Mathematics behind Maximum Margin Classification, Hard Margin SVM, Soft Margin	**18**

Module 3	SVM, Non-Linear SVM, The Kernel Trick, Different Kernel Functions (Linear Kernel, Polynomial Kernel, RBF Kernel, Sigmoid Kernel); K-Nearest Neighbour Classifier (KNN):- Working Principle, Illustrative examples on designing KNN Classifiers; Naive Bayes Classifier - Working Principle, Types of Naive Bayes Classifiers, Illustrative examples. Cross-Validation (k-fold)	
Module 4	Ensemble Learning:- Basic Techniques (Voting, Simple and Weighted Averaging Techniques), Advanced Techniques (Bagging, Boosting, Stacking) - The Random Forest Classifier Assessment of Classification Models:- Performance measures- Confusion Matrix, Accuracy, Precision, Recall, F1- Score, Receiver Operating Characteristics Curve (ROC), Area Under the Curve (AUC), Illustrative examples on Classification model assessment, Class Imbalance techniques	**18**

References

1. Mitchell, T. M. (2017). *Machine learning* (1st ed.). McGraw-Hill Education.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7
3. Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
4. Alpaydin, E. (2015). *Introduction to machine learning* (3rd ed.). MIT Press.

Video Links

https://youtu.be/W0UPlwETZWM?si=OR-iRu9pybEhVYsJ-MLE Estimator

https://youtu.be/qpGo87YV1R0?si=DvOiuRj973ogvhak -MAP Estimator

https://youtu.be/MXTsSXIa4i0?si=QtAToZt5EbtQ6914- Linear Regression

https://youtu.be/d_bynTkE9yY?si=2Q2ZlVxSjdVsQWRO-LogisticRegression

https://youtu.be/yG1nETGyW2E?si=CKpG_MY4PqsQjdd -Decision Trees

https://youtu.be/SRVswRH5Q7E?si=JViz8hye8WIzXjNe-SVM

https://youtu.be/2ATqoglcHus?si=6h1LTNeKsuHWHfwP-KNN

https://youtu.be/uQSn3oLVu-8?si=1XnPc-YZco0HGqz7- Naive Bayes

SPSTAIC03: Machine Learning II

Course Objectives:

1. To introduce the learners the in depth theory of neural networks and their optimization.
2. To enable the learners to design solutions for materials sciences issues using unsupervised learning.

Module	Syllabus Description	Contact hours
Module 1	Introduction to Neural Networks:- Biological Neuron versus Artificial Neuron, The Perceptron Model, Notion of Bias and Activation Functions, Different Activation Functions (Sign Function, Sigmoid Function, Hyperbolic Tangent Function, ReLU Function, Leaky ReLU function, Softmax function). Architecture of Neural Networks:- Feedforward and Feedback Neural Networks, Single Layer Perceptron, Linear Separability, The Perceptron Learning Algorithm and its Convergence,	**18**
Module 2	Multilayer Neural Networks:- Multilayer Perceptron, Representational Capability, Universal Approximation Capability, Kolmogorov's Theorem, Size versus depth, MLP Representational Capability. Back Propagation Algorithm:- Back Propagation Learning Law, The Back Propagation Learning Algorithm, Issues in Back Propagation Learning (Vanishing and Exploding Gradient Problems, Local Minima and Saddle Points, Overfitting, Slow Convergence, Sensitivity to the selection of the Initial weights, Computational Cost), Loss functions	**18**
Module 3	Methods to Rectify the Issues in Backpropagation :- Gradient Clipping, Batch Normalization, Regularization, Dropout, Use of Adaptive Learning Rate Methods (Back Propagation Algorithm with Momentum, Adaptive Gradient Descent Algorithm (ADAGRAD), Root Mean Square Propagation (RMSProp), Adaptive Momentum Estimation (ADAM)) Weight Initialization	**18**

	Techniques:- Random Initialization methods, Non-random Initialization methods	
Module 4	Methods to Rectify the Issues in Backpropagation :- Gradient Clipping, Batch Normalization, Regularization, Dropout, Use of Adaptive Learning Rate Methods (Back Propagation Algorithm with Momentum, Adaptive Gradient Descent Algorithm (ADAGRAD), Principal Component Analysis (PCA), Root Mean Square Propagation (RMSProp), Adaptive Momentum Estimation (ADAM)) Weight Initialization Techniques:- Random Initialization methods, Non-random Initialization methods, Autoencoders-Introduction	**18**

References

1. Reed, R. D., & Marks, R. J. II. (1999). *Neural smithing: Supervised learning in feedforward artificial neural networks*. MIT Press.
2. Aggarwal, C. C. (2018). *Neural networks and deep learning*. Springer.
3. Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python: A guide for data scientists*. O'Reilly Media.
4. Aggarwal, C. C. (2023). *Neural networks and deep learning: A textbook* (2nd ed.). Springer.

Video Links

1. Introduction to Neural Networks
   https://youtu.be/xbYgKoG4x2g?si=6qYK55wDC18JrpV6
2. Perceptron Convergence
   https://youtu.be/tRG-OnnQ9g4?si=za8uY4Tc3Q23Ep4A
3. Back Propagation Algorithm
   https://youtu.be/nz3NYD73H6E?si=XYauYJbTUW2Q8pET
4. Principal Component Analysis
   https://youtu.be/H0HjNuNvFVI?si=Ee_e0VRDLmViDSFo
5. Dimensionality Reduction
   https://youtu.be/HnVYF6VQryU?si=46_e8bO4eiEPevU9
6. Unsupervised Learning
   https://youtu.be/NhimXdFenrg?si=WZoxnlT2zqj77Fwj

SPSTAIC04:Deep Learning Concepts

Course Objectives:

1. To give the learners an understanding about the foundations of deep learning architecture and applications.
2. To equip the learners with the necessary skills to set up neural network architecture and use it for materials science applications.

Module	Syllabus description	Contact hours
Module 1	Review of Neural Networks:- Biological Neuron vs. Artificial Neuron, Activation Functions (Step, Sign, Sigmoid, Softmax, tanh, ReLU, LreLU, PReLU functions), Single Layer Perceptron, Linear Separability, XOR Problem, Multilayer Perceptron, Backpropagation algorithm, Hyperparameter tuning.	**18**
Module 2	Convolutional Neural Networks (CNN):- CNN-Architectural Overview, Motivation, Layers, Filters, Parameter sharing, Regularization, Convolution and Pooling as infinitely strong prior, Variants of Convolution Operation, Efficient Convolution Computation Methods, Popular CNN Architectures (ResNet, Alexnet, VGGNet, Inception), Transfer Learning	**18**
Module 3	Recurrent Neural Networks (RNN):- Model of a typical RNN, Computation Graph, Bidirectional RNNs, Encoder – decoder sequence to sequence architectures, BPTT for training RNN, Deep Recurrent Networks, Recursive Recurrent Networks, Long Short Term Memory Networks (LSTM), Gated Recurrent Networks (GRU), Transformers-Introduction	**18**
Module 4	Unsupervised Deep Learning Models:- AutoEncoders(AE), Types of AutoEncoders (Undercomplete AE, Sparse AE, Deep AE, Contractive AE, Denoising AE, Convolutional AE, Variational AE), Boltzmann machines, Types of Boltzmann machines (Restricted Boltzmann Machines, Deep Belief Networks, Deep Boltzmann Machines) , Introduction - Explainable AI, Model deployment basics, Ethical AI	**18**

References

(1) Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

(2) Nielsen, M. A. (2015). *Neural networks and deep learning*. Determination Press.

(3) Aggarwal, C. C. (2023). *Neural networks and deep learning: A textbook* (2nd ed.). Springer. **DOI:** 10.1007/978-3-031-29642-0

(4) Gibson, A., & Patterson, J. (2017). *Deep learning: A practitioner's approach.* O'Reilly Media.

(5) Chollet, F. (2021). *Deep learning with Python* (2nd ed.). Manning Publications.

Video Links

https://nptel.ac.in/courses/106105215

SPSTAIP01:Machine Learning LAB

Course Objectives:

1. To enable the learners to design and implement solutions based on machine learning techniques for Materials Science problems.
2. To provide the ability to analyse and evaluate various machine learning algorithms.

Expt no:	Experiments (Programming Language: Python)
1	Design and Implementation of Simple Linear Regression –Materials dataset
2	Design and Implementation of Multivariable Linear Regression
3	Design and Implementation of Logistic Regression
4	Design and Implementation of Decision Trees
5	Design and Implementation of Naive Bayes Classifier
6	Design and Implementation of Perceptron models to perform simple arithmetic and logic operations
7	Design and Implementation of the K-means Clustering
8	Design and Implementation of PCA
9	Design and Implementation of Ensembling methods , Bagging and boosting

SPSTAIP02: Deep Learning Lab

Course Objectives:

1. To get hands-on experience in machine learning.
2. To enable the learners to develop deep learning models for Raman Spectroscopic data and XRD pattern using python.

1	Implement and demonstrate Single, Multi variable and Polynomial Regression for a given set of training data stored in a .CSV file and evaluate the accuracy.
2	Write a Python program to implement KNN classifiers and calculate the accuracy, precision, and recall for your data set (Minerals dataset- RRuff dataset)
3	Assuming a set of data that need to be classified, use a Naive Bayes classifier to perform this task and evaluate the accuracy(Minerals dataset- RRuff dataset).
4	Write a Python program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate materials science data set for building the decision tree and apply this knowledge to classify a new sample.
5	Implement hierarchical clustering algorithm on a given dataset to categorize the data.
6	Implement k means clustering algorithm on a given dataset to categorize the data.
7	Build an Artificial Neural Network using Backpropagation algorithm on a given dataset and test the same with appropriate dataset(Minerals dataset- RRuff dataset).
8	Implement Feed forward neural network with three hidden layers for classification on minerals dataset.