Paper code: ARM210	L	P	Credits
Subject: Introduction to Machine Learning	4	0	4

Marking Scheme:
Teachers Continuous Evaluation: As per university examination norms from time to time.
End Term Theory Examination: As per university examination norms from time to time.

INSTRUCTIONS TO PAPER SETTERS: Maximum Marks : AS per University norms

- There should be 9 questions in the end term examination question paper
- Question No. 1 should be compulsory and cover the entire syllabus. This question should have objective or short answer type questions.
- Apart from Question No. 1, the rest of the paper shall consist of four units as per the syllabus. Every unit should have two questions. However, students may be asked to attempt only 1 question from each unit.
- The questions are to be framed keeping in view the learning outcomes of course/paper. The standard/ level of the questions to be asked should be at the level of the prescribed textbooks.
- The requirement of (scientific) calculators/ log-tables/ data-tables may be specified if required

Course Outcomes [Bloom's Knowledge Level (KL)]:
CO1: Ability of students to understand the basics concepts of machine learning and Data Science **[K1, K2]**
CO2: Ability of students to apply and analyze various classification algorithms **[K2,K3,K4]**
CO3: Ability of students to apply and analyze various regression analysis and clustering techniques **[K2,K3,K4]**
CO4: Ability of students to understand the basic concept of deep neural networks and evaluating performance of machine learning algorithms **[K2, K4]**.

CO/PO	PO01	PO02	PO03	PO04	PO05	PO06	PO07	PO08	PO09	PO10	PO11	PO12
CO1	3	3	3	3	3	-	-	-	1	2	1	2
CO2	3	3	3	3	3	-	-	-	2	2	2	2
CO3	3	3	3	3	3	-	-	-	2	2	2	3
CO4	3	3	3	3	3	-	-	-	2	2	2	3

Course Content	**No of lectures**
Unit I **Introduction to Data Science Concepts:** Data Science Terminology, Process, Data Science toolkit, Types of data, Source of Data, data Collection & APIs, Exploring & Fixing Data, Data Storage & Management, Big Data, Big Data Technologies, Stages of Data Analytics. **Introduction to Machine Learning:** Learning Theory, Hypothesis & Target Class, Inductive Bias & Bias Variance Tradeoff, Occam's razor, approximation & estimation error, empirical & expected risk, ERM	**[12]**

Unit II **Supervised learning for classification:** choosing a classification algorithm, implementing perceptron learning rules, modeling class probabilities via logistic regression, maximum margin classification with support vector machine, decision tree learning, k-nearest neighbor algorithm, Bayesian learning, Ensemble learning-majority voting classifier, bagging and boosting classifier, random forest classifier.	**[12]**
Unit III **Predicting continuous target variables with regression analysis:** introducing a simple linear regression model, evaluating the performance of linear regression model, using regularized methods for regression, turning a linear model into a curve - polynomial regression, Non-linear regression model-support vector, decision tree and random forest regressor **Unsupervised Learning:** Clustering-K Means, K Means++, Hierarchical and Density Based, Mixture Models, Expectation Maximization, Non-Parametric Density Estimation.	**[12]**
Unit IV **Evaluation:** Performance evaluation metrics-accuracy score, precision, recall, F1-score, ROC curves, mean squared error, r2-score. Validation methods, Bias variance decomposition, confusion matrix, model complexity. **Introduction to Deep Networks:** Deep Feed Forward networks, convolutional neural networks, stacking, striding and pooling.	**[12]**
Text Books: [T1] Sebastian Raschka,Vahid Mirjalili, (2019), Python Machine Learning - Third Edition, Pact Publisher [T2] Tom M. Mitchell, (1997). Machine Learning, McGraw-Hill [T3] Duda, R. O. & Hart, P. E. (2006). Pattern Classification. John Wiley & Sons.	
Reference Books: [R1] Bishop, C. M. & Nasarabadi, N. M. (2006). Patter Recognition & Machine Learning (Vol. 4, No. 4, P. 738). New York: Springer	