


# Institute of Actuaries of India

## Subject

## CS3 - Data Science: Tools, Techniques and Applications
## Core Principles

## For 2027 Examinations

# Data Science: Tools, Techniques and Applications (CS3)

**Syllabus for 2027 Examinations**

This syllabus provides information to support the study of this subject. It outlines the knowledge and skills required to understand and apply data science and machine learning techniques in actuarial and business contexts.

This syllabus includes:

- Aim of the subject
- Links across the Qualifications
- Topic and topic weightings
- Subject objectives

**Aim**

Data Science: Conceptual Framework, Tools and Techniques (CS3) provides a grounding in the principles and applications of data science and machine learning. The subject develops the knowledge and practical skills required to build, evaluate and apply machine learning models for data-driven decision-making in actuarial and business contexts.

**Links across the Qualifications**

A strong understanding of this subject calls for a reasonable working knowledge of the material covered in:

- CS1 – Actuarial Statistics
- CS2 – Risk Modelling & Survival Analysis

Practical applications of the material covered in this subject are discussed further in:

- CP1 – Actuarial Practice
- All Specialist Principles (SP) subjects
- All Specialist Advanced (SA) subjects

**Topics & topic weightings**

1. The Need for Machine Learning [10%]
2. ML Frameworks and Use Cases in Finance [15%]
3. Feature Engineering [5%]
4. Introduction to Supervised Machine Learning & Regression [10%]

5. Supervised Machine Learning – Classification [5%]
6. Artificial Neural Networks [5%]
7. Time Series Models [5%]
8. Unsupervised Machine Learning [10%]
9. Reinforcement Learning [5%]
10. Ensemble Learning Algorithms [5%]
11. Managing Hyperparameters of ML Models [5%]
12. Managing Hyperparameters of DL Models [5%]
13. Model Evaluation Methods [5%]
14. Machine Learning Pipeline [10%]

**Subject Objectives**

**Topic 1: The Need for Machine Learning [10%]**

1.1 Understand the role of Data Science, Analytics and Machine Learning in supporting data-driven decision-making.

1.2 Describe the key types of analytics and their applications in business and actuarial decision-making.

1.3 Explain the relationship between Artificial Intelligence, Machine Learning and Deep Learning.

1.4 Describe the major forms of Machine Learning and their applications.

1.5 Distinguish between common Machine Learning tasks including Regression, Classification and Clustering.

1.6 Describe common Machine Learning techniques and their applications in business and actuarial contexts.

**Topic 2: ML Frameworks and Use Cases in Finance [15%]**

2.1 Understand the purpose and role of Machine Learning frameworks in managing Machine Learning projects.

2.2 Describe common Machine Learning frameworks including CRISP-DM, KDD, SEMMA and Informs-JTA.

2.3 Explain the phases of the CRISP-DM framework and their application to Machine Learning projects.

2.4 Apply the CRISP-DM framework to Regression and Classification problems.

2.5 Identify Machine Learning applications in finance and actuarial practice.

**Topic 3: Feature Engineering [5%]**

3.1 Understand the concept and importance of Feature Engineering in Machine Learning.

3.2 Apply feature construction techniques for preparing data for Machine Learning models.

3.3 Apply feature extraction techniques for dimensionality reduction and data representation.

3.4 Describe feature selection methods and their role in improving model performance.

3.5 Apply exploratory data analysis techniques to support feature engineering and model development.

**Topic 4: Introduction to Supervised Machine Learning & Regression [10%]**

4.1 Understand the concepts of Supervised Machine Learning and its applications in predictive modelling.

4.2 Distinguish between Regression and Classification problems.

4.3 Apply Regression techniques to develop predictive models and interpret model outputs.

4.4 Understand the assumptions, evaluation and validation of Regression models.

4.5 Apply Decision Tree models for Regression tasks and interpret decision rules.

**Topic 5: Supervised Machine Learning – Classification [5%]**

5.1 Apply Classification techniques for predictive modelling.

5.2 Evaluate the performance of Classification models using appropriate measures.

5.3 Apply Logistic Regression models for classification problems.

5.4 Understand the principles and applications of Naïve Bayes classifiers.

5.5 Apply Decision Tree models for classification tasks and interpret model outputs.

**Topic 6: Artificial Neural Networks [5%]**

6.1 Understand the principles and architecture of Artificial Neural Networks.

6.2 Describe the role of activation functions and loss functions in Neural Networks.

6.3 Apply Deep Neural Networks to Regression and Classification problems.

6.4 Evaluate and validate Neural Network models.

**Topic 7: Time Series Models [5%]**

7.1 Understand the principles of Time Series Analysis and Forecasting.

7.2 Apply Exponential Smoothing techniques for time series forecasting.

7.3 Apply Autoregressive models for time series forecasting.

7.4 Apply Machine Learning techniques to Time Series problems.

**Topic 8: Unsupervised Machine Learning [10%]**

8.1 Understand the principles and applications of Unsupervised Machine Learning.

8.2 Apply clustering techniques including K-Means and Hierarchical Clustering.

8.3 Evaluate clustering solutions using appropriate performance measures.

8.4 Understand the principles and applications of anomaly detection.

8.5 Apply Multivariate Gaussian methods for anomaly detection.

**Topic 9: Reinforcement Learning [5%]**

9.1 Understand the principles and applications of Reinforcement Learning.

9.2 Explain the concepts of Markov Chains and Markov Decision Processes.

9.3 Apply Value Iteration and Policy Iteration techniques to Reinforcement Learning problems.

9.4 Identify business applications of Reinforcement Learning.

**Topic 10: Ensemble Learning Algorithms [5%]**

10.1 Understand the principles and applications of Ensemble Learning.

10.2 Describe Bagging, Boosting, Voting and Stacking methodologies.

10.3 Apply Ensemble Learning techniques to predictive modelling problems.

10.4 Understand the application of Ensemble methods to Time Series and Survival Analysis.

**Topic 11: Managing Hyperparameters of ML Models [5%]**

11.1 Understand the role of parameters and hyperparameters in Machine Learning models.

11.2 Explain the impact of hyperparameters on model performance.

11.3 Apply Grid Search and Random Search techniques for hyperparameter tuning.

11.4 Apply hyperparameter optimisation techniques to Machine Learning models.

**Topic 12: Managing Hyperparameters of DL Models [5%]**

12.1 Understand hyperparameters in Deep Learning models.

12.2 Apply Grid Search, Random Search, Bayesian Optimisation and Hyperband methods.

12.3 Apply hyperparameter optimisation techniques to Deep Neural Networks.

**Topic 13: Model Evaluation Methods [5%]**

13.1 Understand the purpose and importance of model evaluation.

13.2 Apply performance measures for Regression and Classification models.

13.3 Apply validation and cross-validation techniques.

13.4 Explain the concepts of underfitting and overfitting.

13.5 Understand model explainability and feature importance.

**Topic 14: Machine Learning Pipeline [10%]**

14.1 Understand the role of Machine Learning pipelines in model development and deployment.

14.2 Design and implement training pipelines using the CRISP-DM framework.

14.3 Understand the principles of model, feature and hyperparameter deployment.

14.4 Explain the role of serialization and containerization in Machine Learning deployment.

14.5 Understand monitoring, validation, versioning and governance considerations for Machine Learning pipelines.

**END**