SDSC5001: STATISTICAL MACHINE LEARNING I

Effective Term
Semester A 2025/26

Part I Course Overview

Course Title
Statistical Machine Learning I

Subject Code
SDSC - Data Science
Course Number
5001

Academic Unit
Data Science (DS)

College/School
College of Computing (CC)

Course Duration
One Semester

Credit Units
3

Level
P5, P6 - Postgraduate Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
Nil

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract

This course focuses on the theoretical foundation and fundamental methods in statistical machine learning, covering the key concepts of the probability theory and statistical inference for machine learning, classical and cutting-edge methods and theories for regression and classification, and popular methods for unsupervised learning.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Summarize the key concepts of statistical inference theory for machine learning	20	x		
2	Build fundamental statistical regression and classification models	20	x		
3	Formulate kernel machines and regularization forms	20	x	x	
4	Describe the underlying bias-variance tradeoff in each learning method	20	x	x	
5	Apply taught machine learning methods to solve real life data analytics problem via software	20		x	x

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	In lectures, students will learn fundamental theories and principles in statistical machine learning.	1, 2, 3, 4	26 hours/semester
2	Laboratory	In labs, students will learn software of statistical machine learning and apply taught methods to address real-world problems.	2, 3, 4, 5	13 hours/semester

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks ("-" for nil entry)	Allow Use of GenAI?
1	Group Project	2, 3, 5	20	-	Yes
2	Homework assignments	1, 2, 3, 4	30	-	Yes

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Assessment Rubrics (AR)

Assessment Task

Group Project (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

20%

Excellent

(A+, A, A-) High

Good

(B+, B, B-) Significant

Fair

(C+, C, C-) Moderate

Marginal

(D) Basic

Failure

(F) Not even reaching marginal levels

Assessment Task

Homework assignments (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

30%

Excellent

(A+, A, A-) High

Good

(B+, B, B-) Significant

Fair

(C+, C, C-) Moderate

Marginal

(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Examination (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion
50%

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair
(C+, C, C-) Moderate

Marginal
(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Group Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion
20%

Excellent
(A+, A, A-) High

Good
(B+, B) Significant

Marginal
(B-, C+, C) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Homework assignments (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion
30%

Excellent
(A+, A, A-) High

Good
(B+, B) Significant

Marginal
(B-, C+, C) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Examination (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion
50%

Excellent
(A+, A, A-) High

Good
(B+, B) Significant

Marginal
(B-, C+, C) Basic

Failure
(F) Not even reaching marginal levels

Part III Other Information

Keyword Syllabus
- Review of Probability Theory
- Statistical Inference: Method of Moments, Maximum Likelihood Estimation, Bootstrap
- Regression: Linear and Nonlinear Regression, Kernel Smoothing, Local Polynomial, Cubic Splines, Regression Splines, Gaussian Process Regression
- Classification: Misclassification Error, Discriminant Analysis, Logistic Regression, CART, Bagging, Random Forest, Boosting
- Kernel Machines: Loss Functions, Regularization Form, SVM
- Unsupervised Learning: Principle Component Analysis, Clustering

Reading List

Compulsory Readings

	Title
1	Statistical Inference by George Casella and Roger L. Burger
2	The Elements of Statistical Learning by Hastie, Tibshirani and Friedman, Springer
3	Lecture Notes

Additional Readings

	Title
1	Nil