EE3001: FOUNDATIONS OF DATA ENGINEERING

Effective Term
Semester A 2023/24

Part I Course Overview

Course Title
Foundations of Data Engineering

Subject Code
EE - Electrical Engineering
Course Number
3001

Academic Unit
Electrical Engineering (EE)

College/School
College of Engineering (EG)

Course Duration
One Semester

Credit Units
3

Level
B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
MA1200 Calculus and Basic Linear Algebra I, or MA1300 Enhanced Calculus and Linear Algebra I

Precursors
MA1201 Calculus and Basic Linear Algebra II, or MA1301 Enhanced Calculus and Linear Algebra II

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract

The aim of this course is to provide students with an understanding of the basic probability and statistical techniques used in data engineering.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Apply probability techniques for modelling and understanding data		x	x	
2	Apply statistical tools to analyse data and understand their physical meanings and implications		x	x	
3	Acquire a basic understanding of how the techniques learnt can be applied to data engineering applications		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.

Teaching and Learning Activities (TLAs)

	TLAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	Key concepts are explained during lecture.	1, 2, 3	3 hrs/wk

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Tests (min.: 2)	1, 2, 3	30	
2	#Assignments (min.: 3)	1, 2, 3	20	

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Additional Information for ATs

Remark:
To pass the course, students are required to achieve at least 30% in coursework and 30% in the examination.

may include mini projects, in-class assignments, and homework assignments.

Assessment Rubrics (AR)

Assessment Task

Tests (2)

Criterion

Understanding of concepts and ability to apply them successfully to solve engineering problems. The correct choice of probability and statistical models in different scenarios is also tested.

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

Assignments

Criterion

The same abilities in tests at the same level as tests, but allow formative learning through searching information, peer discussion and group learning.

Ability to apply knowledge learnt to daily life applications is also assessed.

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

Criterion
The achievement levels of the CILOs.

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

Part III Other Information

Keyword Syllabus

Introduction
A bird eye's view of the course with an emphasis on probabilistic and statistical techniques and tools for data engineering applications, and the use and possible miss-uses

Probability models for data engineering
Essential concepts in probability, Bayes' reasoning, Markov and Chebyshev's inequality, representation of data and events as random variables, popular probability distributions in data engineering, ways to represent relationships between random variables, sampling methods

Statistical tools and their data engineering applications
Statistical descriptors, parameter estimation, parametric and non-parametric hypothesis testing, correlation, linear regression and model building, prediction, useful engineering tools and software, assumptions, precautions and limits of statistics, common pitfalls in interpreting data

Reading List

Compulsory Readings

	Title
1	S.M. Ross, Introduction to Probability and Statistics for Engineers and Scientists, 5th Ed., 2014 (e-book).

Additional Readings

	Title
1	D. Wackerly, W. Mendenhall, and R.L. Scheaffer, Mathematical statistics with applications, 7th Ed., 2007.
2	N.J. Salkind, Statistics for people who hate statistics, 4th Ed., 2017.
3	R.E. Walpole et al., Probability and statistics for engineers and scientists, 9th Ed., 2014.