

GUJARAT TECHNOLOGICAL UNIVERSITY (GTU)
Competency-focused Outcome-based Green Curriculum-2021 (COGC-2021)

Semester –V
 Course Title: AI and ML
 (Course Code:4353205)

Diploma Programme in which this course is offered	Semester in which offered
Information & Communication Technology	5 th semester

1. RATIONALE

Artificial Intelligence (AI) and Machine Learning (ML) deals with the making of intelligent machines and computer programs. It is concerned with building smart machines capable of performing tasks that typically require human intelligence. Machine learning is a sub-part of Artificial Intelligence that focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. Machine learning focuses on developing computer programs that can access data and use it to learn for themselves. In the modern world, knowing AI and ML provides students a competitive advantage. AI and ML basic and recent trends are covered in this course. The larger area of artificial intelligence (AI) includes ANN and NLP as essential components. In order to tackle problems in the real world, this aids students in developing fundamental neural network models.

2. COMPETENCY

The purpose of this course is to help the student to attain the following industry identified competency through various teaching-learning experiences:

- Utilize AI and ML algorithms using python programming to solve well defined problems.

3. COURSE OUTCOMES (COs)

The practical exercises, the underpinning knowledge, and the relevant soft skills associated with this competency are to be developed in the student to display the following COs:

- a) Understand fundamental principles of Artificial Intelligence.
- b) Compare types of machine learning.
- c) Build a simple Neural Network model to solve real world problem.
- d) Apply data preprocessing on text/paragraph using NLTK library.
- e) Extend with recent advancement in AI and ML field.

4. TEACHING AND EXAMINATION SCHEME

Teaching Scheme (In Hours)			Total Credits (L+T/2+P/2)	Examination Scheme				Total Marks
				Theory Marks		Practical Marks		
L	T	P	C	CA	ESE	CA	ESE	
3	0	2	4	30	70	25	25	150

(*): ^{AI and ML} Out of 30 marks under the theory CA, 10 marks are for assessment of the micro-project to facilitate integration of COs and the remaining 20 marks is the average of 2 tests to be taken during the semester for the assessing the attainment of the cognitive domain UOs required for the attainment of the COs.

Legends: *L*-Lecture; *T* – Tutorial/Teacher Guided Theory Practice; *P* -Practical; *C* – Credit, *CA* - Continuous Assessment; *ESE* -End Semester Examination.

5. SUGGESTED PRACTICAL EXERCISES

The following practical outcomes (PrOs) are the subcomponents of the COs. These PrOs need to be attained to achieve the COs.

Sr. No.	Practical Outcomes (PrOs)	Unit No.	Approx. Hrs.required
1	a) Introduction to Pandas for data import and export (Excel, CSV etc.) b) Basic Introduction to Scikit learn c) Numerical Computing & Analytics with Python (NumPy, Matplotlib)	I	04
2	Build a machine learning model that can learn from the measurements of iris flower whose species is known, so that we can predict the species for a new iris flower. Use Iris dataset, a classical dataset in machine learning and statistics included in <i>scikit-learn</i> in the <i>datasets</i> module of python. Use a k-nearest neighbors classifier to Build this model.	II	04
3	Implement following activation functions using python to build simple neural network. a. ReLU b. Sigmoid c. Tanh	III	04
4	Implement following feed forward neural network using python programming: a. Single layer feed forward neural network. b. Multi-layer feed forward neural network.	III	04
5	Perform following data preprocessing on text/paragraph using NLTK library: a. Write a Python program to tokenize words, sentence wise. b. Write a python program that accepts the list of tokenized word and stems it into root word. c. Write a program in python to identify the part of speech for each word in the text. d. Write a Python NLTK program to remove stop words from a given text. e. Write a python program for identifying and correcting misspelled words in a given text, such as an essay or a letter.	IV	04
6	Implement following Word embedding techniques in NLP. a. TFIDF- Term Frequency Inverse document Frequency b. BOW (Bag of Words) c. Word2Vec	IV	04
7	Demonstrate Text generation using BERT model.	V	02
8	Compare results of different prompt queries on the same topic using LLM.	V	02
	Total		28

Note

- i. More **Practical Exercises** can be designed and offered by the respective course teacher to develop the industry relevant skills/outcomes to match the COs. The above table is only a suggestive list.
- ii. The following are some **sample** 'Process' and 'Product' related skills (more may be added/deleted depending on the course) that occur in the above listed **Practical Exercises** of this course required which are embedded in the COs and ultimately the competency..

S. No.	Sample Performance Indicators for the PrOs	Weightage in %
1	Prepare the dataset	25
2	Relevant implementation of AI models using suitable libraries or frameworks.	20
3	Fulfillment of specific requirements for the given task and appropriate handling of domain-specific challenges.	25
4	Use of relevant evaluation metrics for the given task.	10
5	Applying best practices and well-structured documentation.	20
Total		100

6. MAJOR EQUIPMENT/ INSTRUMENTS REQUIRED

This major equipment with broad specifications for the PrOs is a guide to procure them by the administrators to usher in uniformity of practical in all institutions across the state.

S. No.	Equipment Name with Broad Specifications	PrO. No.
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM, Python versions: 2.7.X, 3.6.X	All
2	Python IDEs and Code Editors Open Source : Anaconda Navigator	

7. AFFECTIVE DOMAIN OUTCOMES

The following **sample** Affective Domain Outcomes (ADOs) are embedded in many of the above- mentioned COs and PrOs. More could be added to fulfill the development of this competency.

- a) Work as a data scientist.
- b) Follow ethical practices.

The ADOs are best developed through the laboratory/field based exercises. Moreover, the level of achievement of the ADOs according to Krathwohl's 'Affective Domain Taxonomy' should gradually increase as planned below:

- i. 'Valuing Level' in 1st year
- ii. 'Organization Level' in 2nd year.
- iii. 'Characterization Level' in 3rd year.

8. UNDERPINNING THEORY

Only the major Underpinning Theory is formulated as higher-level UOs of *Revised Bloom's taxonomy* in order development of the COs and competency is not missed out by the students and teachers. If required, more such higher-level UOs could be included by the course teacher to focus on the attainment of COs and competency.

Unit	Unit Outcomes (UOs) (4 to 6 UOs at Application and above level)	Topics and Sub-topics
Unit – I Introduction of Artificial Intelligence (AI)	1a. Describe basic concept of Artificial Intelligence 1b. Differentiate between AI and ML and area of AI 1c. Identify various applications of Artificial Intelligence 1d. Introduction of existing machine learning python libraries.	1.1 Definition of AI, Future of Artificial Intelligence, History and Types of AI, AI ethics and limitations 1.2 Difference between AI & ML Artificial Intelligence Areas like Expert Systems, Natural Language Processing, Neural Networks, Robotics, Fuzzy Logic Systems 1.3 AI applications in various industries (healthcare, finance, manufacturing, etc.) 1.4 Pandas, Numpy, Matplotlib, Scikit-learn, 1.4.1 Different visualization (graphs like bar, line, pie, scatter, histogram, box, violin, heatmap) using matplotlib, 1.4.2 Different Data Analytics (Descriptive, Exploratory, inferential, predictive, prescriptive) with hypothetical/fictional/real life data using Numpy, pandas etc.
Unit – II Core concepts of Machine learning (ML)	2a. Recognize well-posed machine learning problems. 2b. Explain examples of the various machine learning models. 2c. Explain basic concepts of Reinforcement learning. 2d. Compare types of Machine learning.	2.1 Identify Well-Posed Machine learning Problem 2.2 Types of Machine Learning Models: Supervised, Unsupervised, Reinforcement 2.3 Reinforcement Learning: - Terms, Key features - Approaches to implement reinforcement learning: Value based, Policy Based, Model based, - Elements of reinforcement learning- Policy, reward signal, value function, Model; - Types of reinforcement learning: Positive and Negative Comparison 2.4 Comparison between Supervised, unsupervised and reinforcement learning with example

Unit– III Neural Network Fundamentals	3a.Understand fundamentals of Neural networks. 3b.Implement activation functions to build neural network. 3c.Implement types of ANN. 3d.Describe weight learning process in ANN.	3.1 Introduction to Neural Networks: Understanding the Biological Neuron, Exploring the Artificial Neuron 3.2 Types of Activation Functions: ReLU, Sigmoid, Hyperbolic Tangent Function 3.3 Architectures of Neural Network:
		Single-layer feed forward network, Multi-layer feed forward ANNs, Recurrent network 3.4 Learning Process in ANN: Number of layers, Direction of signal flow, Number of nodes in layers, Weight of interconnection
Unit– IV Foundation of Natural Language processing (NLP)	4a.Understand basics of NLP. 4b.Discuss components of NLP. 4c.Explain phases of NLP. 4d.Apply data preprocessing techniques on text/ paragraph/ document.	4.1 Introduction To NLP: History of NLP, Advantages of NLP, Disadvantages of NLP 4.2 Components of NLP : Natural Language Understanding (NLU) ,Natural Language Generation (NLG) 4.3 Phases of NLP: Lexical Analysis, Syntactic Analysis, Semantic Analysis, Discourse Integration, Pragmatic Analysis 4.4 Data Preprocessing Using NLTK: Tokenization, Frequency Distribution of Words, Filtering Stop Words, Stemming, Lemmatization, Parts Of Speech (POS) Tagging, Name Entity Recognition, WordNet 4.5 Types of Ambiguities in NLP
Unit– V Advance Topics in AI and ML	5a.Understand the concept and applications of Automated Machine Learning (AutoML). 5b.Describe the fundamentals and applications of Generative Adversarial Networks (GANs). 5c.Explain the architecture and usage of the BERT model in Natural Language Processing (NLP). 5d.Explore the development and impact of Large Language Models (LLMs).	5.1 Automated Machine Learning (AutoML): <ul style="list-style-type: none"> • Overview of AutoML, • Key Components of AutoML • Applications of AutoML 5.2 Generative Adversarial Networks (GANs): <ul style="list-style-type: none"> • Fundamentals of GANs • Components of GANs • Applications of GANs 5.3 BERT Model in Natural Language Processing (NLP): <ul style="list-style-type: none"> • Overview of the BERT Model • Architecture of BERT • Pre-training and Fine-tuning • Applications of BERT 5.4 Large Language Models (LLMs):

		<ul style="list-style-type: none"> • Introduction to Prompt engg. • Basic introduction of LLM. • Capabilities of models like GPT-3 and GPT-4. • Applications of LLMs
--	--	--

Note: The UOs need to be formulated at the 'Application Level' and above of Revised Bloom's Taxonomy' to accelerate the attainment of the COs and the competency.

9. SUGGESTED SPECIFICATION TABLE FOR QUESTION PAPER DESIGN

Unit No.	Unit Title	Teaching Hours	Distribution of Theory Marks			
			R Level	U Level	A	Total Marks
I	Introduction of Artificial Intelligence (AI)	07	4	6	4	14
II	Core concepts of Machine learning (ML)	09	4	8	-	12
III	Neural Network Fundamentals	09	4	6	6	16
IV	Foundation of Natural Language processing	09	4	6	6	16
V	Advance Topics in AI and ML	08	4	8	-	12
Total		42	20	34	16	70

Legends: R=Remember, U=Understand, A=Apply and above (Revised Bloom's taxonomy)

Note: This specification table provides general guidelines to assist students for their learning and to teachers to teach and question paper designers/setters to formulate test items/questions assess the attainment of the UOs. The actual distribution of marks at different taxonomy levels (of R, U and A) in the question paper may vary slightly from the above table.

10. SUGGESTED STUDENT ACTIVITIES

Other than the classroom and laboratory learning, following are the suggested student-related *co-curricular* activities which can be undertaken to accelerate the attainment of the various outcomes in this course: Students should conduct following activities in group and prepare reports of about 5 pages for each activity, also collect/record physical evidences for their (student's) portfolio which will be useful for their placement interviews:

- Explore different data repositories and register for ML based competitions on platforms like kaggle.
- Undertake micro-projects in teams.
- Give a seminar on any relevant topics.
- Collect data from snapchat/ facebook and do sentiment analysis using NLP.

11. SUGGESTED SPECIAL INSTRUCTIONAL STRATEGIES (if any)

These are sample strategies, which the teacher can use to accelerate the attainment of the various outcomes in this course:

- Massive open online courses (*MOOCs*) may be used to teach various topics/subtopics.
- Guide student(s) in undertaking micro-projects.
- 'L' in section No. 4 means different types of teaching methods that are to be employed by teachers to develop the outcomes.
- About **20% of the topics/sub-topics** which are relatively simpler or descriptive in nature

is to be given to the students for *self-learning*, but to be assessed using different assessment methods.

- e) With respect to *section No.11*, teachers need to ensure to create opportunities and provisions for *co-curricular activities*.
- f) Guide students for open source python editors.

12. SUGGESTED MICRO-PROJECTS

Only one micro-project is planned to be undertaken by a student that needs to be assigned to him/her in the beginning of the semester. In the first four semesters, the micro-project are group-based. However, in the fifth and sixth semesters, it should be preferably be *individually* undertaken to build up the skill and confidence in every student to become problem solver so that s/he contributes to the projects of the industry. In special situations where groups have to be formed for micro-projects, the number of students in the group should *not exceed three*.

The micro-project could be industry application based, internet-based, workshop- based, laboratory-based or field-based. Each micro-project should encompass two or more COs which are in fact, an integration of PrOs, UOs and ADOs. Each student will have to maintain a dated work diary consisting of individual contributions in the project work and give a seminar presentation of it before submission. The total duration of the micro-project should not be less than *16 (sixteen) student engagement hours* during the course. The student ought to submit a micro-project by the end of the semester to develop the industry oriented COs.

A suggestive list of micro-projects is given here. This has to match the competency and the COs. Similar micro-projects could be added by the concerned course teacher:

- **Project idea 1: Music Genre Classification:** Develop a classifier to categorize music clips into different genres (e.g., rock, jazz, hip-hop) using audio features extracted from the clips.
- **Project idea 2: Sentiment Analysis on Movie Reviews:** Use a dataset of movie reviews labeled with sentiment (positive/negative) to build a classifier that predicts the sentiment of movie reviews.

Algorithms to explore: Naive Bayes, Logistic Regression, Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks.

- **Project idea 3: Predicting Student Performance:** Build a model to predict student performance (e.g., pass/fail, grade) based on features such as study time, previous grades, and demographics.
- **Project idea 4: Conversational Bots:** ChatBots, most tech companies today use Chatbots, which are conversational bots, to communicate with their consumers and handle issues. It is an excellent method for both customers and businesses to save time.

13. SUGGESTED LEARNING RESOURCES

S.No.	Title of Book	Author	Publication with place, year and ISBN
1	Introduction to Machine Learning with Python	Andreas C. Müller and Sarah Guido	O'Reilly, 2016
2	Machine Learning_ Step-by-Step Guide To Implement Machine Learning Algorithms with Python.	Rudolph Russell	Rudolph Russell Publications
3	Machine Learning	Saikat Dull, S.Chjandramouli	Das, Pearson

4	Natural Language Processing with Python	Steven Bird, Ewan Klein & Edward Loper	O'Reilly, 2009
---	---	--	----------------

SOFTWARE/LEARNING WEBSITES

- <https://www.geeksforgeeks.org/machine-learning/>
- <https://www.geeksforgeeks.org/natural-language-processing-nlp-tutorial/>
- https://www.tutorialspoint.com/machine_learning_with_python/index.htm
- <https://onlinecourses.nptel.ac.in/>
- <https://www.kaggle.com/learn>

14. PO-COMPETENCY-CO MAPPING

Semester V	AI and ML (Course Code:)						
	POs and PSOs						
Competency & Course Outcomes	PO 1 Basic & Discipline specific knowledge	PO 2 Problem Analysis	PO 3 Design/development of solutions	PO 4 Engineering Tools, Experimentation & Testing	PO 5 Engineering practices for society, sustainability & environment	PO 6 Project Management	PO 7 Life-long learning
Competency Applying Artificial Intelligence and Machine Learning concepts to solve real world problems.							
Course Outcomes							
CO a) Understand fundamental principles of Artificial Intelligence.	3	-	-	-	-	-	-
CO b) Compare types of machine learning.	3	2	2	2	2	2	2
CO c) Build a simple Neural Network model to solve real world problem.	3	2	3	3	3	3	2
CO d) Apply data preprocessing on text/paragraph using NLTK library.	3	2	2	3	2	2	2
CO e) Extend with recent advancement in AI and ML field.	3	2	3	3	3	3	3

Legend: '3' for high, '2' for medium, '1' for low or '-' for the relevant correlation of each competency, CO, with PO/ PSO

15. COURSE CURRICULUM DEVELOPMENT

COMMITTEE GTU Resource Persons

Sr. No	Name and Designation	Institute	Email
1	M. R. Makwana Lecturer-EC	L.E. College(Diploma)-Morbi	mahesh.ecengg@gmail.com