COURSEPACK

The scheme is an overview of work-integrated learning opportunities and gets students outinto the real world. This will give what a course entails.

Course Title	Machine Learning			Course Type	Comprehensive	
Course Code	E2UC508C			Class	B. Tech.	

Instruction delivery	Activity	Credits	Credit Hours	Total Number of Classes per Semester				Assessmentin Weightage	
				Theory	Tutorial	Practical	Self-study	CIE	SEE
	Lecture	3	3						
	Tutorial	0	0						
	Practical	0	0						
	Self-study	1	1						
	Total	**4**	**4**	**45**	**0**	**0**	**15**	**50%**	**50%**

Names Course Instructors	Course Lead:	Dr. Santosh Kumar	
	Theory		**Practical**

Theory
1 Santosh Kumar
2 Abhist Kumar
3 Akhilesh Kumar Singh (32563)
4 Akhilesh Kumar Tripathi
5 C. Ramesh Kumar
6 Vartika Mishra
7 Ashish Ranjan
8
9 Kaleemur Rehman
10 Arunendra Mani Tripathi
11
12 Vijayant Pawar
13 Gopal Chandra Jana
14 Abdul Mazid
15 Anupam Kumar Sharma
16 Krishna Kant Agrawal
17 Shivani Goswami
18 Sugan Patel
19
20 G. Sakthi
21 Geetika Sharma
22 Rakesh Bharti
23 Mili Dhar
24 Albert Mundu

COURSE OVERVIEW

Machine Learning (ML) gives computers the ability to learn without being explicitly programmed. Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people. Machine learning allow for

computers to train on data inputs and use statistical analysis in order to output values that fall within a specific range. Machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on data inputs.

PREREQUISITE COURSE

PREREQUISITE COURSE REQUIRED	YES	NO (√)

If, yes please fill in the details

Prerequisite: Basic knowledge of Databases and Python

Prerequisite course code	Prerequisite course name

COURSE OBJECTIVE

- To understand concepts of machine learning systems, types of learning etc.
- To apply the concepts of Machine learning for predictive analysis and evaluate the models.
- To create machine learning models and analyze the performance.

COURSE OUTCOMES(COs)

After the completion of the course, the student will be able to:

CO No.	Course Outcomes
E2UC508C.1	apply the concepts of machine learning including different types of learning to solve real world problems.
E2UC508C.2	implement the ML models like classification, regression, and clustering algorithms and evaluate the models on parameters like accuracy, Precision recall, etc.
E2UC508C.3	analyze the performance of ML models and conduct comparative analysis.
E2UC508C.4	create probabilistic and unsupervised learning models to handle unknown patterns.

CO numbering method: **CO   2   01   1**
(Semester no/ course no/ co sequence)

BLOOM'S LEVEL OF THE COURSE OUTCOMES

Comprehensive

CO No.	Remember BTL1	Understand BTL2	Apply BTL3	Analyse BTL4	Evaluate BTL5	Create BTL6
E2UC508C.1						

ACK |

E2UC508C.2			∫		∫	
E2UC508C.3				∫		
E2UC508C.4						∫

PROGRAM OUTCOMES (POs):

PO1 Computing Science knowledge: Apply the knowledge of mathematics, statistics, computing science and information science fundamentals to the solution of complex computer application problems.

PO2 Problem analysis: Identify, formulate, review research literature, and analyze complex computing science problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and computer sciences.

PO3 Design/development of solutions: Design solutions for complex computing problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4 Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5 Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern computing science and IT tools including prediction and modeling to complex computing activities with an understanding of the limitations.

PO6 IT specialist and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional computing science and information science practice.

PO7 Environment and sustainability: Understand the impact of the professional computing science solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8 Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the computing science practice.

PO9 Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10 Communication: Communicate effectively on complex engineering activities with the IT analyst community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

Assignment, Quiz, Class test, SWAYAM/NPTEL/MOOCs and etc.

PO11 Project management and finance: Demonstrate knowledge and understanding of the computing science and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12 Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAMME SPECIFIC OUTCOME (PSO):

The students of Computer Science and Engineering shall:

PSO1: Have the ability to work with emerging technologies in computing requisite to Industry 4.0.

PSO2: Demonstrate Engineering Practice learned through industry internship and research project to solve live problems in various domains.

COURSE ARTICULATIONMATRIX

COs#/ POs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
E2UC508 C.1	2	2	2		2								2	2
E2UC508 C.2	2	2	2		1								1	1
E2UC508 C.3	2	2	2		1								2	1
E2UC508 C.4	1	1	1	1	2								2	1

Note: 1-Low, 2-Medium, 3-High

COURSE ASSESSMENT

Assessment Pattern for Comprehensive Course:

Type of Course (C)	CIE			Total Marks		Final Marks CIE*0.5+SEE*0.5
	LAB@ (Work+ Record)	MTE	Course-based Project^	CIE	SEE	
COMPREHENSIVE	25	50	25	100	100	100

@Lab Work-15 marks + Lab Record-10 marks

COURSE CONTENT

THEORY

Content
Basic concepts: Definition of learning systems, Goals and applications of machine learning. Aspects of developing a learning system: training data, concept representation, function approximation. History of ML, Introduction of Machine Learning Approaches – (Artificial Neural Network, Clustering, Reinforcement Learning, Decision Tree Learning, Bayesian networks, Support Vector Machine, Genetic Algorithm), Issues in Machine Learning and Data Science Vs Machine Learning.   **Types of Learning:** Supervised learning and unsupervised learning. Overview of classification: setup, training, test, validation dataset, over fitting.   **Classification:** Decision Trees – Attribute Selection Measures and Tree Pruning; Bayesian and Rule-based Classification; Model Evaluation and Selection; Cross-Validation; Classification Accuracy; Bayesian Belief Networks; Classification by Backpropagation; and Support Vector Machine   **Regression:** Linear Regression and Logistic Regression, Perceptron, Exponential family, Generative learning algorithms, Gaussian discriminant analysis, Naive Bayes, Support vector machines: Optimal hyper plane, Kernels. Model selection and feature selection. Combining classifiers: Bagging, boosting (The Ada boost algorithm), Evaluating and debugging learning algorithms, Classification errors. Introduction to Weka: a machine learning tool for Classifiers Introduction to Weka: a machine learning tool for Clustering   **Decision Tree Learning** - Decision tree learning algorithm, Inductive bias, Inductive inference with decision trees, Entropy and information theory, Information gain, ID-3 Algorithm, Issues in Decision tree learning. INSTANCE-BASED LEARNING – k-Nearest Neighbour Learning, Locally Weighted Regression, Radial basis function networks, Case-based learning   **Clustering**, K-means, EM Algorithm, Mixture of Gaussians, Factor analysis, PCA (Principal components analysis), ICA (Independent components analysis), latent semantic indexing, Spectral clustering, Markov models Hidden Markov models (HMMs).   **Reinforcement Learning**–Introduction to Reinforcement Learning , Learning Task,Example of Reinforcement Learning in Practice, Learning Models for Reinforcement – (Markov Decision process , Q Learning - Q Learning function, Q Learning Algorithm), Application of Reinforcement Learning, Introduction to Deep Q Learning.   **Genetic Algorithms:** Introduction, Components, GA cycle of reproduction, Crossover, Mutation, Genetic Programming, Models of Evolution and Learning, Applications.

PRACTICAL

Sl. No.	Content
1	The probability that it is Friday and that a student is absent is 3%. Since there are 5 school days in a week, the probability that it is Friday is 20%. What is the probability that a student is absent given that today is Friday? Apply Baye's rule in python to get the result. (Ans: 15%)
2	Extract the data from database using python
3	Implement k-nearest neighbours classification using python
4	Given the following data, which specify classifications for nine ombinations of VAR1 and VAR2 predict a classification for a case where VAR1=0.906 and VAR2=0.606, using the result of k-means clustering with 3 means (i.e., 3 centroids)
5	The following training examples map descriptions of individuals onto high, medium and low credit-worthiness. Input attributes are (from left to right) income, recreation, job, status, age-group, home-owner. Find the unconditional probability of 'golf' and the conditional probability of 'single' given 'medRisk' in the dataset
6	Implement linear regression using python
7	Implement naive baye's theorem to classify the English text
8	Implement an algorithm to demonstrate the significance of genetic algorithm
9	Implement the finite words classification system using Back-propagation algorithm
10	Given the following vectors: A = [1, 2, 3, 4, 5, 6, 7, 8, 9 10] B = [4, 8, 12, 16, 20, 24, 28, 32, 36, 40] C = [10, 9, 8, 7, 6, 5, 4, 3, 2, 1] Ex. 1: Find the arithmetic mean of vector A, B and C Ex. 2: Find the variance of the vector A, B and C Ex. 3: Find the Euclidean distance between vector A and B Ex. 4: Find the correlation between vectors A & B and A & C
11	Load breast cancer dataset and perform classification using Euclidean distance. Use 70% data as training and 30% for testing.
12	To install the pypi libraries for Machine Learning.
13	Review of python datatypes for Artificial Intelligence and Machine Learning
14	Develop a machine learning model for standard database using Support Vector Machines
15	Develop a machine learning model for standard database using Decision Trees.
16	Develop a machine learning model for standard database using Random Forest.
17	Deploy standardization and normalization on some standard dataset.
18	Deploy Principal Component Analysis to extract relevant features on some standard database.
19	Understanding of github and conda environments.

20	To use the github packages and libraries to frame a standard project and commit back to github.
21	Implement and demonstrate the "find-s" algorithm for finding the most specific hypothesis based on a given set of training data samples: read the training data from csv file.
22	For a given set of training data example stored in .csv file, implement and demonstrate the Candidate-Elimination Algorithm to output and describes the set of all hypotheses consistent with training example.
23	Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
24	Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets
25	Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets
26	Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set
27	Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Java/Python ML library classes/API.
28	Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. Add Python ML library classes/API in the program
29	Write a program to implement K-nearest neighbour algorithm to classify iris dataset. Print both correct and wrong predication using python machine learning
30	Implement the non-parametric Locally Weighted Regression algorithm to fit data points. Select appropriate data set for your experiment and draw graphs.

LNo	Topic for Delivery	Skill	Competency
1	Basic concepts: Definition of learning systems, Goals and applications of machine learning	Student will be able to understand basic concepts of ML	
2	Aspects of developing a learning system: training data	Student will be able to understand training Data Set	Student will be able to understand and apply ML concepts on real data set
3	concept representation, function approximation	Student will be able to understand concept representation	
4	History of ML, Introduction of Machine Learning Approaches – (Artificial Neural Network, Clustering)	Student will be able to understand history of ML	
5	History of ML, Introduction of Machine Learning Approaches – (Reinforcement Learning, Decision Tree Learning)		
6	History of ML, Introduction of Machine Learning Approaches – (Bayesian networks, Support Vector Machine, Genetic Algorithm)		
7	Issues in Machine Learning and Data Science Vs Machine Learning	Student will be able to understand difference between DS and ML	
8	Types of Learning: Supervised learning and unsupervised learning	Student will be able to understand types of learning	
9	Overview of classification: setup, training, test	Student will be able to understand classification of objects	Student will be able to understand, apply and build models for classification and prediction based on real data set
10	Overview of classification: validation dataset, over fitting		
11	Classification Families: linear discriminative, non-linear discriminative, decision trees		
12	Classification: Decision Trees – Attribute Selection Measures and Tree Pruning;		
13	Bayesian and Rule-based Classification; Model Evaluation and Selection; Cross-Validation;	Student will be able to understand and apply classification models	
14	Classification Accuracy; Bayesian Belief Networks; Classification by Backpropagation; and Support Vector Machine		
15	Regression: Linear Regression and Logistic Regression, Perceptron, Exponential family		
16	Generative learning algorithms, Gaussian discriminant analysis, Naive Bayes,	Student will be able to understand and implement classifiers and	
17	Support vector machines: Optimal hyper plane, Kernels. Model selection and feature selection		

18	Combining classifiers: Bagging, boosting (The Ada boost algorithm)	regression models	Student will be able to understand ML tool and use the tool for predictive analysis and performance measurement of the models
19	Combining classifiers: Evaluating and debugging learning algorithms, Classification errors		
20	Introduction to Weka: a machine learning tool for Classifiers	Student will be able to use weka ML tool	
21	Introduction to Weka: a machine learning tool for Clustering		
22	Decision Tree Learning - Decision tree learning algorithm	Student will be able to understand and use decision tree	
23	Inductive bias, Inductive inference with decision trees		
24	Entropy and information theory, Information gain	Student will be able to understand, apply and build models based on entropy	
25	ID-3 Algorithm		
26	Issues in Decision tree learning		
27	INSTANCE-BASED LEARNING – k-Nearest Neighbour Learning	Student will be able to understand, apply instance and case based learning	
28	Locally Weighted Regression, Radial basis function networks		Student will be able to understand and apply some advanced concepts for classification, prediction and analysis
29	Case-based learning		
30	Clustering, K-means	Student will be able to understand, apply and build models based clustering	
31	EM Algorithm	Student will be able to understand and apply PCA, ICA and HMM	
32	Mixture of Gaussians, Factor analysis		
33	PCA (Principal components analysis)		
34	ICA (Independent components analysis)		
35	latent semantic indexing, Spectral clustering		
36	Markov models Hidden Markov models (HMMs).		
37	Reinforcement Learning–Introduction to Reinforcement Learning	Student will be able to understand, apply and build models based on Reinforcement	
38	Learning Task,Example of Reinforcement Learning in Practice		

39	Learning Models for Reinforcement – (Markov Decision process, Q Learning - Q Learning function, Q Learning Algorithm)	Learning	
40	Application of Reinforcement Learning		
41	Introduction to Deep Learning	Student will be able to understand and apply Deep learning and GA	Student will be able to understand and apply deep learning and evolutionary approach for ML
42	Genetic Algorithms: Introduction, Components		
43	GA cycle of reproduction, Crossover, Mutation		
44	Genetic Programming		
45	Models of Evolution and Learning, Applications		

LESSON PLAN FOR COMPREHENSIVE COURSES

FOR THEORY   15 weeks * 3 Hours = 45 Classes) (1credit = 1Lecture Hour)
FOR PRACTICAL 15 weeks * 2Hours = 30 Hours lab sessions (1 credit = 2 lab hours)

BIBLIOGRAPHY

- **Text Book**

1. Tom M. Mitchell, ―Machine Learning, McGraw-Hill Education (India) Private Limited, 2013.

2. Andreas C. Müller & Sarah Guido, "Introduction to Machine Learning", with Python, A GUIDE FOR DATA SCIENTISTS, O'Reilly.

3. Stephen Marsland, ―Machine Learning: An Algorithmic Perspective, CRC Press, 2009.

4. Nils J. Nilsson, "INTRODUCTION TO MACHINE LEARNING", TEXTBOOK, Robotics Laboratory Stanford University, Stanford, CA 94305, 1998

5. M. Gopal, "Applied Machine Learning", McGraw Hill Education

6. Richard S. Sutton and Andrew G. Barto, "Reinforcement Learning: An Introduction", A Bradford Book The MIT Press, Cambridge, Massachusetts, London, England

- **Reference Books**

1. Chris Albon, Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning, O'Reilly, 2018

2. Probability for Statistics and Machine Learning: Anirban DasGupta - 2011

3. Rudolph Russell, "Machine Learning: Step-by-Step Guide To Implement Machine Learning Algorithms with Python"

4. Peter Harrington, Machine Learning in Action , April 2012 ISBN 9781617290183

5. David Barber, "Bayesian Reasoning and Machine Learning", 2012

- Journals/Magazines/Govt. Reports/Gazatte/Industry Trends
- **Webliography**
  1. https://ifacet.iitk.ac.in/professional-certificate-course-in-ai-and-machine-learning/?utm_source=google&utm_medium=cpc&utm_term=machine%20learning%20program&utm_content=20161035334-154966829171-661014543750&utm_device=c&utm_campaign=Search-DataCluster-PG-DSAI-AIML-CAIML-IITK-Main-IITKDomain-IN-AllDevice-Core-adgroup-CAIML-IITK-Core-Keywords&gclid=CjwKCAjwv8qkBhAnEiwAkY-ahtf1DaJ-Y1LwSBrj-txj8SC4XbNNvivTroiHn11CXXzGox61khUlNhoCtXAQAvD_BwE
  2. https://www.udemy.com/topic/machine-learning/
  3. https://mrcet.com/downloads/digital_notes/CSE/IV%20Year/MACHINE%20LEARNING(R17A0534).pdf
  4. https://www.cs.cmu.edu/~hn1/documents/machine-learning/notes.pdf
  5. https://news.vidyaacademy.ac.in/wp-content/uploads/2018/10/NotesOnMachineLearningForBTech-1.pdf

- **SWAYAM/NPTEL/MOOCs Certification** (Two-Three numbers)
  1. https://www.udemy.com/topic/machine-learning/
  2. https://www.coursera.org/learn/machine-learning
  3. https://developers.google.com/machine-learning/crash-course
  4. https://www.edx.org/learn/machine-learning

PROBLEM-BASED LEARNING

Exercises in Problem-based Learning (Assignments)

SNo	Problem	BTL
1	Autonomous Driving: Develop a machine learning model to enable self-driving cars to navigate safely and make real-time decisions on the road.	BTL6
2	Cancer Diagnosis: Build a model to accurately detect and classify different types of cancer cells from medical images, aiding in early diagnosis and treatment.	BTL6
3	Natural Language Processing: Create a language model that can understand and generate human-like text, enabling applications like chatbots and language translation.	BTL6
4	Credit Card Fraud Detection: Develop a model to identify fraudulent credit card transactions and prevent financial losses for both cardholders and banks.	BTL6
5	Sentiment Analysis: Build a model that can analyze and classify sentiment in social media posts or customer reviews, helping businesses understand customer feedback.	BTL6
6	Recommendation Systems: Create a personalized recommendation system that suggests products, movies, or music based on user preferences and behavior.	BTL6
7	Stock Market Prediction: Develop a model that can predict stock market trends and make informed investment decisions.	BTL6
8	Disease Outbreak Prediction: Build a model to predict and track the spread of infectious	BTL6

	diseases, aiding in early detection and prevention.	
9	Customer Churn Prediction: Develop a model to identify customers who are likely to churn or cancel their subscriptions, allowing businesses to take proactive measures to retain them.	BTL6
10	Speech Recognition: Create a model that can accurately transcribe spoken language, enabling applications like voice assistants and transcription services.	BTL6
11	Fraudulent Insurance Claims Detection: Build a model to identify fraudulent insurance claims, reducing losses for insurance companies and ensuring fair payouts.	BTL6
12	Energy Consumption Forecasting: Develop a model to predict energy consumption patterns, helping utility companies optimize energy production and distribution.	BTL6
13	Object Detection in Images: Create a model that can detect and classify objects in images, enabling applications like autonomous robots and surveillance systems.	BTL6
14	Human Activity Recognition: Build a model that can recognize and classify human activities from sensor data, aiding in healthcare monitoring and fitness tracking.	BTL6
15	Spam Email Detection: Develop a model that can accurately identify and filter out spam emails, improving email security and user experience.	BTL6
16	Document Classification: Create a model that can automatically classify documents into different categories, aiding in information retrieval and organization.	BTL6
17	Facial Recognition: Build a model that can recognize and identify individuals from facial images, enabling applications like biometric authentication and surveillance.	BTL6
18	Customer Segmentation: Develop a model that can segment customers into distinct groups based on their behavior and preferences, aiding in targeted marketing campaigns.	BTL6
19	Traffic Flow Prediction: Build a model to predict traffic congestion and optimize traffic flow, improving transportation efficiency and reducing travel time.	BTL6
20	Anomaly Detection: Create a model that can detect unusual patterns or outliers in data, aiding in fraud detection, network security, and system monitoring.	BTL6
21	Predictive Maintenance: Develop a model that can predict equipment failures or maintenance needs, enabling proactive maintenance and reducing downtime.	BTL6
22	Air Quality Prediction: Build a model to predict air pollution levels and provide early warnings, aiding in public health and environmental monitoring.	BTL6
23	Customer Lifetime Value Prediction: Create a model that can estimate the future value of a customer, helping businesses prioritize customer acquisition and retention strategies.	BTL6
24	Gesture Recognition: Develop a model that can recognize and interpret hand gestures, enabling applications like sign language translation and virtual reality interactions.	BTL6
25	Document Summarization: Build a model that can automatically generate concise summaries of long documents, aiding in information extraction and comprehension.	BTL6
26	Fraud Detection in Online Transactions: Create a model to identify fraudulent activities in online transactions, ensuring secure and trustworthy e-commerce experiences.	BTL6
27	Predictive Healthcare: Develop a model that can predict the likelihood of disease or health complications for individuals, aiding in personalized healthcare interventions.	BTL6
28	Social Network Analysis: Build a model to analyze social network data and identify influential users or communities, aiding in targeted marketing and social network security.	BTL6
29	Weather Forecasting: Create a model that can predict weather patterns and provide accurate forecasts, aiding in disaster preparedness and resource planning.	BTL6

30	Customer Satisfaction Prediction: Develop a model that can predict customer satisfaction levels based on various factors, helping businesses improve their products and services.	BTL6
31	Video Surveillance: Build a model that can analyze video feeds and detect suspicious activities or objects, aiding in public safety and security.	BTL6
32	Machine Translation: Create a model that can automatically translate text from one language to another, enabling seamless communication across different languages.	BTL6
33	Medical Image Segmentation: Develop a model that can accurately segment and identify different structures or abnormalities in medical images, aiding in diagnosis and treatment planning.	BTL6
34	Predictive Maintenance in Manufacturing: Build a model that can predict equipment failures or maintenance needs in manufacturing processes, reducing downtime and optimizing production.	BTL6
35	Customer Demand Forecasting: Create a model that can predict customer demand for products or services, aiding in inventory management and supply chain optimization.	BTL6
36	Emotion Recognition: Develop a model that can recognize and classify human emotions from facial expressions or speech, enabling applications like affective computing and mental health monitoring.	BTL6
37	Cybersecurity: Build a model that can detect and prevent cyber threats, such as malware, phishing attacks, or network intrusions, ensuring data security and privacy.	BTL6
38	Autonomous Drones: Create a model that enables drones to navigate autonomously, avoiding obstacles and performing tasks like package delivery or aerial surveillance.	BTL6
39	Financial Market Trading: Develop a model that can analyze market data and make automated trading decisions, optimizing investment strategies and maximizing returns.	BTL6
40	Customer Attrition Prediction: Build a model that can predict customer attrition or churn for subscription-based services, allowing businesses to implement retention strategies.	BTL6
41	Food Quality Inspection: Create a model that can inspect and classify food products based on quality attributes, ensuring food safety and compliance with standards.	BTL6
42	Document Fraud Detection: Develop a model that can detect forged or altered documents, aiding in identity verification and fraud prevention.	BTL6
43	Wildlife Conservation: Build a model that can identify and track endangered species from camera trap images or acoustic data, aiding in wildlife conservation efforts.	BTL6
44	Social Media Influencer Detection: Create a model that can identify influential users or social media influencers based on their engagement and impact, aiding in influencer marketing campaigns.	BTL6
45	Customer Complaint Analysis: Develop a model that can analyze customer complaints and extract valuable insights, helping businesses improve their products and services.	BTL6
46	Energy Load Forecasting: Build a model that can predict energy consumption patterns and forecast future load demands, aiding in energy resource planning and grid management.	BTL6
47	Autonomous Robots: Create a model that enables robots to perform complex tasks autonomously, such as object manipulation, navigation, or human-robot interaction.	BTL6
48	Medical Diagnosis Support: Develop a model that can assist doctors in diagnosing diseases by analyzing patient symptoms, medical records, and test results.	BTL6
49	Social Media Trend Analysis: Build a model that can analyze social media trends and predict viral content or emerging topics, aiding in marketing and content creation strategies.	BTL6
50	Quality Control in Manufacturing: Create a model that can detect defects or anomalies in	BTL6

	manufacturing processes, ensuring product quality and reducing waste.	

STUDENT-CENTEREDLEARNING (SELF-LEARNING TOWARDS LIFE-LONG-LEARNING)

Self-Learning (it's a typical course-based project to be carried out by a whole class in groups of fourstudents each; they should exhibit higher level BTLs)

The students, ina group, are expected toconceive anidea based onthe content(objectives/outcomes) and apply the suitable knowledge todemonstrate their learning.

A list of 30-40 project statements can be offered to the students to choose or students can conceive their ownideas(teamwork),design and develop the product/process/serviceandimplementthesame.

1. Age and Gender Detection with Python

Facial analysis from images has gained a lot of interest because it helps in several different problems like better ad targeting for customers, better content recommendation system, security surveillance, and other fields as well. Age and gender are a very important part of facial attributes and identifying them is the very basic of facial analysis and a required step for such tasks. Many companies are using these kinds of tools for different purposes making it easier for them to work with customers, cater to their needs better, and create a great experience for them. It is easier to identify and predict the needs of people based on their gender and age.

2. Amazon Alexa Reviews Sentiment Analysis

Amazon Alexa is a cloud-based voice service developed by Amazon that allows customers to interact with technology. There are currently over 40 million Alexa users around the world, so analyzing user sentiments about Alexa will be a good data science project. So, if you want to learn how to analyze the sentiments of users using Amazon Alexa, this article is for you. The machine learning project of Amazon Alexa Reviews Sentiment Analysis Using Python can be a good option.

3. Amazon Product Reviews Sentiment Analysis

Amazon is an American multinational corporation that focuses on e-commerce, cloud computing, digital streaming, and artificial intelligence products. But it is mainly known for its e-commerce platform which is one of the biggest online shopping platforms today. There are so many customers buying products from Amazon that today Amazon earns an average of $ 638.1 million per day. So, for having such a large customer base, it will turn out to be an amazing data science project if we can analyze the sentiments of Amazon product reviews. So, the Amazon Product Reviews Sentiment Analysis project with Python can be the best option for you.

4. Amazon Recommendation System

Recommendation Systems are one of the widely used applications of Data Science in most companies based on products and online services. Amazon is a great example of such companies. Being an online shopping website, Amazon needs to generate personalized recommendations to provide a better user experience. The Recommendation System of Amazon follows the principle of generating product-based recommendations which means you get to measure the similarities between two products and recommend the most similar products to the user. The methods of measuring similarities between two products have always been a major focus of researchers.

5. Autocorrect Keyboard with Python and Machine Learning

Almost every smartphone brand irrespective of its price provides an autocorrect feature on their keyboards. In the context of machine learning, autocorrect is based on natural language processing. As the name suggests it is programmed to correct spellings and

errors while typing. The Autocorrect model is programmed to correct spellings and errors while inputting text and locating the most comparable and related words. It is completely based on NLP that compares the words in the vocabulary dictionary and the typed words on the keyboard. If the typed word is found in the dictionary, the autocorrect feature assumes you typed the correct term. If the word does not exist, the tool identifies the most comparable words from the smartphone's history.

6. Automatic License Number Plate Recognition System

This project aims to recognize license number plates. In order to detect license number plates, you will be using OpenCV to identify number plates and python pytesseract to extract characters and digits from the number plates. OpenCV is an open-source machine learning library and provides a common infrastructure for computer vision. Pytesseract is a Tesseract-OCR Engine to read image types and extract the information present in the image.

7. Automatic Time Series Forecasting

Automatic Time Series Forecasting is a forecast of future values generated over time from past data. Think of how the price changes every day for your favorite stock. Time-series forecasting is can predict the price of that stock over multiple time periods. For example, forecasting what Tesla's stock price will be for the next 60 days or across other time periods. Other examples of time-series data are the weekly numbers of account signups, daily revenue, hourly transactions, and so on.

8. Barbie with Brains Project

The rapid growth of modern-day technology has paved way for innovative ideas, one of which is presented in this paper, which is "Barbie with Brains". This Barbie is contradictory to the other dolls which stay idle, and perhaps interact with humans especially kids, just like any typical person would do. This interactive Barbie becomes more charismatic with its breathtaking features, like Barbie itself being a knowledge hub for education purposes, which benefits children in their schooling and learning, where sometimes there is no need for any knowledge or teaching backup, while Barbie is around. Some of its fascinating physiognomy makes kids feel comfortable with their own toys by initiating conversation, recognizing faces, detecting emotions, and playing comforting songs and relatable messages.

9. Build a Collaborative Filtering Recommender System in Python

Collaborative Filtering is the most common technique used when it comes to building intelligent recommender systems that can learn to give better recommendations as more user information is collected. Collaborative filtering is a technique that can filter out items that a user might like on the basis of reactions by similar users. It searches among a large group of people and finds a smaller set of users with tastes similar to a particular user. It looks at the items they like and combines them to create a ranked list of suggestions. There are many ways to decide which users are similar and combine their choices to create a list of recommendations.

10. Build a Similar Images Finder ML Project

In this project, you will build a similar image finder by dissecting the trained weights of

the image object-classifier VGG and using it to extract feature vectors from an image database to see which images are "similar" to each other. This technique is called transfer learning and requires no training on your end as the hard work was done back in the day when VGG was actually being trained, you can just re-use the trained weights to build a new model.

11. Build Classification Algorithms for Digital Transformation [Banking]

Bank XYZ has a growing customer base where the majority of them are liability customers (depositors) vs. borrowers (asset customers). The bank is interested in expanding the borrower's base rapidly to bring in more business via loan interests. A campaign that the bank ran in the last quarter showed an average single-digit conversion rate. Digital transformation is the core strength of the business strategy – devising effective campaigns with better target marketing to increase the conversion ratio to double-digit with the same budget as per the last campaign. As a data scientist, you are asked to develop a machine learning model to identify potential borrowers to support focused marketing. Build a machine learning model to perform focused digital marketing by predicting the potential customers who will convert from liability customers to asset customers.

12. Build CNN for Image Colorization using Deep Transfer Learning

Image colorization is the process of taking an input grayscale (black and white) image and then producing an output colorized image that represents the semantic colors and tones of the input. In image colorization, a color is assigned to each pixel of a target grayscale image. Image colorization technique is helpful for developing many applications such as Medical Microscope, Medical Imagery, Denoising and recreating old Images, Night Vision Camera, etc. For example, you can use a fully automated data-driven technique called autoencoders for image colorization. Autoencoders are a specific type of feedforward neural network where the input is the same as the output. The VGG16 model will be used as a feature extractor. VGG16 is a classic neural network used as a backbone for many computer vision tasks. This project aims to build a convolutional neural network that will best convert the grayscale images to RGB images.

13. Cartoonify Image with Machine Learning

In this Project web app project, you can directly select an image and then convert it into a cartoon. It's a very interesting project. This is a simple and basic level project highly recommended for learning purposes. Also, you can modify this system as per your requirements and develop a perfect advanced-level project.

Required modules in this project are- CV2 (imported to use OpenCV for image processing), and easy gui (imported to open a file box). It allows us to select any file from our system, Numpy(Images are stored and processed as numbers). These are taken as arrays. We use NumPy to deal with arrays, Imageio (used to read the file which is chosen by the file box using a path), Matplotlib (a library used for visualization and plotting), and OS (is imported to form the plot of images) for OS interaction. Here, to read the path and save images to that path, Flask (a micro web framework written in Python) is used.

14. Catching Illegal Fishing Project

Overfishing and illegal fishing are becoming big problems around the world. For example, there are many records of intensive and often illegal fishing in West African waters by Asian and European fleets that reduce the regular catch for the local populations, increasing their poverty levels. "Being able to see which vessels are fishing where would be a tremendous help in reducing illegal fishing," says Josephus Mamie, head of Sierra Leone's Fisheries Research Unit. In this project, you will collaborate with Global Fishing Watch to detect fishing activity in the ocean using data from the satellite Automatic Identification System (AIS) collected from different vessels around the world. The AIS data contains the latitude, longitude, speed, and course of the vessels at different times.

15. Census Income Dataset Project

A census is the process of collecting, compiling, and publishing demographic, economic, and social data pertaining to a specific time to all persons in a country or delimited part of a country. As part of a census count, most countries also include a census of housing. It is the process of collecting, compiling, and publishing information on buildings, living quarters, and building-related facilities such as sewage systems, bathrooms, and electricity, to name a few. In this project, we will use a standard imbalanced machine learning dataset referred to as the "Adult Income" or simply the "adult" dataset. For census salary prediction, we have to classify salaries that fall within a specified range. This mainly helps to understand the real estate demands and also, the demands for basic amenities based on one's salary range.

16. Classification Projects on Machine Learning for Beginners

In machine learning, Classification is one of the most widely used techniques with various applications. For sentiment analysis, spam detection, risk assessment, churn prediction, and medical diagnosis classification have served as very simple yet powerful methods. In this project, we aim to give you hands-on experience and theoretical explanations of various ensemble techniques. Understanding various Ensemble techniques and implementing them to predict license status for the given business. The dataset used is a licensed dataset. It contains information about 86K different businesses over various features. The target variable is the status of the license, which has five different categories.

17. Classification with Neural Networks

Autoencoders are the simplest of the deep learning architectures. They are a specific type of feedforward neural network where the input is first compressed into a lower-dimensional code. Then, the output is reconstructed from the compact code representation or summary. Therefore, autoencoders have three components built inside them – encoder, code, and decoder. To begin the development process, you will need an encoding method, a decoding method, and a loss function. Binary cross-entropy and mean squared error are the two top choices for the loss function. And to train the autoencoders, you can follow the same procedure as artificial neural networks via back-propagation. Now, let us discuss the applications of these networks.

18. Count Objects in Image

Counting objects in an image is a task of computer vision. There are many computer vision libraries that you can use for this task, such as OpenCV, TensorFlow, PyTorch, Scikit-image, and cvlib. You must have not heard much about the cvlib library in Python. Well, this is a very simple, high-level, and easy-to-use computer vision library in Python. By using the features of this library, we can count the number of objects in an image using Python. To use this library, make sure you have OpenCV and TensorFlow installed in your systems. You can easily install it by using the pip command; pip installs cvlib.

19. Coupon Purchase Prediction

Recruit Ponpare is Japan's leading joint coupon site, offering huge discounts on everything from hot yoga to gourmet sushi, to a summer concert bonanza. Ponpare's coupons open doors for customers they've only dreamed of stepping through. They can learn difficult to acquire skills, go on unheard-of adventures, and dine like (and with) the stars.

Using past purchases and browsing behavior, this competition asks you to predict which coupons a customer will buy in a given period of time. The resulting models will be used to improve Ponpare's recommendation system, so they can make sure their customers don't miss out on their next favorite thing.

20. Covid-19 Detection

The current COVID-19 pandemic threatens human life, health, and productivity. AI plays an essential role in COVID-19 case classification as we can apply machine learning models on COVID-19 case data to predict infectious cases and recovery rates using chest x-ray. Accessing patients' private data violates patient privacy and the traditional machine learning model requires accessing or transferring whole data to train the model. In recent years, there has been increasing interest in federated machine learning, as it provides an effective solution for data privacy, centralized computation, and high computation power.

21. Currency Exchange Rate Prediction

Predicting the currency exchange rates is the regression problem in machine learning. There are changes in exchange rates every day that affect the income of a person, or a business and can even affect the economy of a country. Thus, predicting the currency exchange rates can help an individual as well as a country in many ways. There are so many machine learning algorithms that we can use to predict future currency exchange rates. You can also use artificial neural networks for this task.

22. Customer Churn Prediction Analysis

A well-known bank has been observing a lot of customers closing their accounts or switching to competitor banks over the past couple of quarters. This has caused a huge dent in their quarterly revenues and might drastically affect annual revenues for the ongoing financial year, causing stocks to plunge and market cap to reduce significantly. The idea is to be able to predict which customers are going to churn so that necessary actions/interventions can be taken by the bank to retain such customers.

In this machine learning churn prediction project, we are provided with customer data

pertaining to his past transactions with the bank and some demographic information. We use this to establish relations/associations between data features and customers' propensity to churn and build a classification model to predict whether the customer will leave the bank or not. We also go about explaining model predictions through multiple visualizations and give insight into which factor(s) are responsible for the churn of the customers.

This project walks you through a complete end-to-end cycle of a data science project in the banking industry, right from the deliberations during the formation of the problem statement to making the model deployment-ready.

23. Deep Learning Project for Text Detection in Images using Python

Text detection is the process of detecting the text present in the image. Several applications include solving the captcha, identifying vehicles by reading their license plates, etc. Convolutional neural networks are deep learning algorithms that are very powerful for the analysis of images. On the other hand, Recurrent Neural Networks (RNNs) are used for sequential data such as text. RNNs are ideal for solving problems where the sequence is more important than the individual items themselves. This model is best used for images with a single line of text in them so we will build this model on images with single-line texts. This project aims to build a convolutional recurrent neural network that can detect the text from a given image.

24. Deepfake Detection with Machine Learning

Machine learning algorithms such as a generative adversarial network (GAN) can be used to create deepfakes. Discriminative models can be used as a method of detecting deepfake videos. Generative adversarial networks (GANs) are an approach to training generative models, in which two neural nets work together to generate fake images that look real and have never been seen before. The first network is called the "generator" and it creates new fakes. The second network is called the "discriminator," which tries to detect whether the images are real or fake. The discriminative model can be used as a method of detecting deepfake videos by using adversarial learning techniques for example, where an attacker's system trains itself on examples of deepfake videos in order to fool a detector.

25. Dog and Cat Classification

The main goal of this tutorial is to develop a system that can identify images of cats and dogs. The input image will be analyzed and then the output is predicted. The model that is implemented can be extended to a website or any mobile device as per the need. The dataset contains a set of images of cats and dogs. Our main aim here is for the model to learn various distinctive features of cats and dogs. Once the training of the model is done it will be able to differentiate images of cat and dog. Adaptive Moment Estimation (Adam) is a method used for computing individual learning rates for each parameter. For the loss function, we are using Binary cross-entropy to compare the class output to each of the predicted probabilities. Then it calculates the penalization score based on the total distance from the expected value.

26. Dogecoin Price Prediction

Predicting the price of a cryptocurrency is a regression problem in machine learning. Bitcoin is one of the most successful examples of cryptocurrency, but we recently saw a major drop in bitcoin prices due to dogecoin. Unlike bitcoin, dogecoin is very cheap right now, but financial experts are predicting that we may see a major increase in dogecoin prices.

There are many machine learning approaches that we can use for the task of Dogecoin price prediction. You can train a machine learning model or you can also use an already available powerful model like the Facebook Prophet Model. But in the section below, you will be using the auto library in Python for the task of Dogecoin price prediction with machine learning.

27. Driver Demand Prediction

Food delivery supported through advanced applications has emerged as one of the fastest-growing developments in the e-commerce space. We all love to order online, one thing that we don't like to experience is variable pricing for delivery charges. Delivery charges highly depend on the availability of riders in your area, the demand of orders in your area, and the distance covered. Due to driver unavailability, there is a surge in delivery pricing and many customers drop off resulting in loss to the company. To tackle such issues if we track the number of hours a particular delivery executive is active, we can efficiently allocate certain drivers to a particular area depending on demand.

28. Emojify – Create your own emoji with Python

Emojis or avatars are ways to indicate nonverbal cues. These cues have become an essential part of online chatting, product review, brand emotion, and many more. It also led to increasing data science research dedicated to emoji-driven storytelling.

With advancements in computer vision and Machine learning, it is now possible to detect human emotions from images. In this deep learning project, we will classify human facial expressions to filter and map corresponding emojis or avatars.

29. End-to-end Fake News Detection System

The end-to-end fake news detection system is one of the top machine learning projects for students or aspiring tech professionals to work on. This machine learning project uses the applications of natural language processing approaches to detect fake news efficiently and effectively. The model can be developed on the basis of the count vectorizer as well as a TFIDF matrix (Term Frequency Inverse Document Frequency). Thus, developers can assemble a dataset consisting of real news and fake news for the model to successfully use the end-to-end fake news detection system.

30. End-to-end Machine Learning Project

An end-to-end machine learning project makes a machine learning portfolio valuable. It can be successfully completed by using a supervised learning regression problem with access to a large dataset. One can also use any dataset from e reliable and safe repository to download and load the file into a pandas datagram. Developers must check for the data type of columns, null values, outliers in the horsepower column, and many more.

31. End-to-end Spam Detection System

This machine learning project is a must-try for aspiring tech professionals with sufficient knowledge of the detection of spam alerts with machine learning. One can use the streamlet library in Python to develop this end-to-end spam detection system. It is essential to write a command mentioned below in the command prompt or terminal to run the code efficiently and effectively.

32. Enron Investigation Project

Techies can start creating impressive machine learning portfolios with the Enron investigation project. The target is to remain focused on developing a Person of Interest (PoI) identifier as well as building a machine learning algorithm for predicting the possible PoI based on multiple different features. One can use scikit-learn as well as machine learning methodologies with features from multiple data. PoIs can include a settlement, plea deal with a government, testifying in exchange for a prosecution summary, and many more.

33. Text Classification with Transformers-RoBERTa and XLNet Model

The text classification with transformers like RoBERTa and XLNet model machine learning project helps to have a deep understanding of how to load, fine-tune, as well as evaluate multiple transformer models for different text classification tasks. The data here needs to be put in tsv format with four columns including guide, lable, alpha, and text, and without any header.

34. Fake Currency Detection with Machine Learning

Fake currency is roaming around in the financial sector more often than before. Thus, one needs to know how to use machine learning for fake currency detection. It is known as a task of binary classification in machine learning to differentiate between fake and real currency all the time without any failure. The dataset usually consists of four input characteristics such as the variance of the image transformed into wavelets, asymmetry of the image transformed into wavelets, Kurtosis of the image transformed into wavelets, as well as image entropy.

35. Fake News Detection

This machine learning project helps to explore different functions of machine learning through news detection approaches for diversified datasets. This fake news detection project needs to cover a wide range of formats as well as subjects to choose from. Relevant data is used to evaluate different views, global events, and many more to extract valuable post features on multiple social media platforms.

36. FEAST Feature Store Example for Scaling Machine Learning

FEAST feature store example for scaling machine learning is one of the important machine learning projects for students and aspiring tech professionals. To complete this project, one needs to set up FEAST repository, understand the entities, feature view, and architecture, and how to retrieve data efficiently and effectively. There are uses of random forest model training, gradient boosting model training, and many more to create this project successfully.

37. Fraud Detection with Machine Learning

Fraud detection with machine learning models and algorithms is useful for tech professionals to work with for a better hands-on experience. It also includes the use of supervised learning, unsupervised learning, semi-supervised learning, as well as reinforcement learning. This project of effective fraud detection with machine learning can become a key tool to prevent cybercrimes.

38. Gender Classification

Gender classification is one of the top machine learning projects in this tech-driven market. it is known as a very difficult task with a real-time application based on the face recognition system. One can use Bag of Words, Sale Invariant Fourier Transform, K-means clustering, and others for effective feature extractors as well as classification. There are picture preparing strategies that can be used to naturally recognize faces, genders, and many more information.

39. Gender Detection

Gender detection is gaining popularity in the list of machine learning projects for the increase in the applications of social media platforms. There are two main features in social attributes through facial attributes — gender and age. Developers can build a real-time gender detection system with machine learning and deep learning efficiently. The model can detect males and females after detecting faces through a convolutional neural network. There are three convolutional layers in this gender detection project with 96 nodes, 256 nodes, and 384 nodes.

40. Gold Price Prediction

For the gold price prediction machine learning project, one needs to use machine learning regression techniques for predicting the accurate price of one of the most valuable and flourishing metals across the world, gold. The model is known for gaining relevant information from the past Gold ETF prices and providing accurate gold price predictions for the very next day. The data may consist of the daily Gold ETF price for the last 12 years to predict the Gold ETF close price.

41. Google Play Store Sentiment Analysis

Google Play Store sentiment analysis project uses both machine learning and Python. The sentiment analysis task is to analyze any customer's reviews and comments by downloading relevant datasets from Kaggle. It is needed to add three new columns in the dataset for a better understanding of the sentiments of each customer review categorized as positive, negative, as well as neutral.

42. Handwriting Recognition

Handwriting recognition is one of the top machine learning projects that are very useful for different purposes in the future. The machine learning algorithm is known for performing handwriting recognition and can recognize characteristics from different media such as images, and touch-screen devices, and convert them to a machine-readable format. There are three categories of the character recognition algorithms such as image pre-processing, feature extraction, as well as classification.

43. Hate Speech Detection

Developers must build one hate speech detection machine learning project with the

integration of Python-based NLP machine learning techniques. The NLP technique is known as Tf-Idf vectorization for extracting relevant keywords that are popular for conveying the importance of hate speech. Logistic regression helps to train computers to classify hate speech with the data extracted from any library or repository.

44. Heart Disease Prediction

This machine learning project is highly useful for the healthcare sector across the world for predicting heart disease to save the life efficiently. The project on predicting a heart disease consists of multiple machine learning algorithms such as neighbors classifier, decision tree classifier, support vector classifier, as well as random forest classifier. One can use multiple libraries for a better understanding of data and different numbers of algorithms to vary their multiple parameters to compare the final model efficiently and effectively.

45. Hotel Reviews Sentiment Analysis

Developers can use sentiment analysis and machine learning algorithms to classify hotel reviews provided by customers from top leading travel sites. There are multiple techniques such as Naïve Bayes, support vector machine, logistic regression, and many more with an ensemble learning model to combine five classifiers and the result will be appropriate. One can collect necessary hotel reviews sentiment analysis from Kaggle and many other places for gaining data such as hotel services during vacations, and business trips.

46. House Price Prediction

The house price prediction with machine learning is one of the key end-to-end projects with the use of advanced regression techniques from Kaggle. The process includes large datasets, cleaning, and pre-processing datasets, fitting a model to the dataset, as well as testing the performance of the model with multiple evaluation metrics. One needs to create a new virtual environment through commands in the terminal. Developers can use a random forest regression algorithm for predicting the house price accurately.

47. Human Activity Recognition using Smartphone Dataset

Human activity recognition(HAR) has wide applications in medical research and human survey system. In this project, we design a robust activity recognition system based on a smartphone. The system uses a 3-D smartphone accelerometer as the only sensor to collect time-series signals. This work focuses on the recognition of human activity using smartphone sensors using different machine learning classification approaches. Data retrieved from smartphones accelerometer and gyroscope sensors are classified to recognize human activity. Results of the approaches used are compared in terms of efficiency and precision.

48. Inventory Demand Forecasting

Inventory forecasting is also known as demand planning, is the practice of using past data and trends. Inventory demand forecasting is the process of predicting customer demand for an inventory item over a defined period. Accurate inventory demand forecasting enables a company to hold the right amount of stock, without over or under-stocking, for optimum inventory control. This project works by helping

companies strike a balance between having too much cash tied up in inventory and having enough stock to meet demand.

49. Language Classification Model

Language classification is the grouping of associated languages in the same category. There are two main kinds of language classification: genealogical and typological classification. Languages are grouped diachronically into language families. language classification is solved here is the classification of text into three possible languages like English, Dutch, and Afrikaans. The goal is to devise a machine-learning algorithm to analyze short conversations extracted from your Project corpora, and automatically classify them according to the language of the conversation.

50. Loan Eligibility Prediction

A loan eligibility prediction using several ML algorithms. The dataset with features, namely, gender, marital status, education, number of dependents, employment status, income, co-applicants income, loan amount, loan tenure, credit history, existing loan status, and property area, are used for determining the loan eligibility regarding the loan sanctioning process. This project uses SQL and Python to build a predictive model on GCP to determine whether an application requesting a loan is eligible or not. This application is working properly and meeting all Banker requirements.

51. Machine Learning in 5 lines of code

This machine learning model is that it is exposed to a large number of inputs and also supplied the output applicable to them. On analyzing more and more data, it tries to figure out the relationship between input and the result. MindsDB is one of the examples of those Machine Learning libraries that are making machine learning easy. By using the MindsDB library we can create a Machine Learning model in under 5 lines of code.

52. Market Basket Analysis

This project is a data mining technique used by retailers to increase sales by better understanding customer purchasing patterns. It involves analyzing large data sets, such as purchase history, to reveal product groupings, as well as products that are likely to be purchased together. This goal is to understand consumer behavior by identifying relationships between the items that people buy. This is a process that looks for relationships among entities and objects that frequently appear together, such as the collection of items in a shopper's cart.

53. MLOps Project for a Mask R-CNN on GCP using uWSGI Flask

MLOps advocates automation and monitoring at all steps of the ML system. It refers to Machine learning operations that represent different methodologies, techniques, and procedures used to automate the deployment and handling of machine learning algorithms. This project is to provide hands-on experience in MLOps by using cloud computing. Google cloud platform is used as a cloud provider. We would advise you to have a basic understanding of Image Segmentation using Mask R-CNN with Tensorflow before jumping into this project.

54. MNIST Handwritten Digit Classification

It is a standard dataset used in computer vision and deep learning. The MNIST dataset is an acronym that stands for the Modified National Institute of Standards and Technology dataset. Although the dataset is effectively solved, it can be used as the basis for learning and practicing how to develop, evaluate, and use convolutional deep learning neural networks for image classification from scratch. MNIST is a dataset of 70,000 images, all images are labeled with the respective digit they represent. MNIST is the hello world of machine learning.

55. Mobile Price Classification

A model based on the Multilayer Perceptron Topology was developed and trained, using a data set, its title is Mobile Price Classification, and was obtained from Kaggle online community, and it is created by Abhishek Sharma. Mobile phones are the best-selling electronic devices as people keep updating their cell phones whenever they find new features in a new device. Introducing a machine learning project on a mobile price classification model to classify the price range of mobiles using Python.

56. Movie Recommendation System using Machine Learning

Recommendation systems are pretty common these days. Netflix, Prime Video, YouTube, and other streaming platforms use these recommendation systems to suggest a movie that you might like to watch according to your previous watch history. A movie recommendation system is an ML-based approach to filtering or predicting the users' film preferences based on their past choices and behavior. It's an advanced filtration mechanism that predicts the possible movie choices of the concerned user and their preferences towards a domain-specific item.

57. Music Recommendation System ML Project

This project will build a music recommendation system using real datasets. It will utilize a dataset sourced from outside called Million Songs Dataset which contains two files: triplet_file and metadata_file. The triplet_file has within it information regarding the singer or band's name, the title of the song, and how long the song lasts. It can actually understand the musical patterns of a listener with their playlist as the source and what factors are really useful in determining the taste and interest of the listener.

58. Netflix Stock Price Prediction

Netflix is one of the most popular OTT streaming platforms. It offers a vast collection of television series and films and owns its productions known as Netflix Originals. To predict the stock prices of Netflix with machine learning, using the LSTM neural network as it is one of the best approaches for regression analysis and time series forecasting. People who are highly active in stock market investments always keep an eye on companies like Netflix because of its popularity. Machine learning has significant applications in stock price prediction.

59. Network Security Analysis

It is an independent evaluation of the informational security of network infrastructure and the preparation of recommendations on raising the security level of the network infrastructure for the best international practices of providing informational security. Network security is the general practice of protecting computer networks and devices

accessible to the network against malicious intent, misuse, and denial. This project is an improvement over traditional network intrusion detection. This dataset is widely used by security data science professionals to classify problems of Network Security.

60. Next Word Prediction Model

It is the task of predicting what word comes next. It is one of the fundamental tasks of NLP and has many applications. Google also uses the next word prediction model based on our browsing history. Google uses our browsing history to make next-word predictions, smartphones, and all the keyboards that are trained to predict the next word are trained using some data. Machine Learning model for next word prediction using Python.

61. NLP-based ChatBot

NLP enables the computer to acquire meaning from inputs given by users. It is a branch of informatics, mathematical linguistics, ML, and AI. This project is a computer program or artificial intelligence that communicates with a customer via textual or sound methods. Such programs are often designed to support clients on websites or via phone. The chatbots are generally used in messaging applications like Slack, Facebook Messenger, or Telegram. NLP-based chatbot project allows users to make chatbots by themself. This is a popular solution for those who do not require complex and sophisticated technical solutions.

62. Object Detection

This project takes an image as input and produces one or more bounding boxes with the class label attached to each bounding box. These algorithms are capable enough to deal with multi-class classification and localization as well as to deal with the objects with multiple occurrences. In object detection, the bounding boxes are always rectangular. Object detection is a fascinating field in machine learning, it is used for research purposes. Some applications of object detection are facial recognition, this can also be used to count people for crowd statistics, also used to identify products, or check the quality of a product.

63. Ola Bike Ride Request Demand Forecast

The Ola service industry is growing for the last couple of years and it is expected to grow in near future. This project is about Ola drivers' need to choose where to hang tight for passengers as they can get somebody quickly. Passengers also prefer a quick taxi service whenever needed. So many times people faced problems with taxi booking requests, which sometimes cannot be fulfilled or the wait time for ride arrival is very long due to the unavailability of a nearby Ola.

64. OpenAI Gym in Machine Learning

It is a toolkit that provides a wide variety of simulated environments. OpenAI is an artificial intelligence research company, funded in part by Elon Musk. OpenAI gym for developing and comparing Reinforcement Learning algorithms. Reinforcement learning is an area of machine learning that allows an intelligent agent to learn the best behaviors in an environment by trial and error. The user can choose their robot, environment, action, and rewards for testing their reinforcement learning algorithms in

OpenAI Gym.

65. OpenCV Project for Beginners to Learn Computer Vision Basics

Single and Multi-Object Tracking is a proper OpenCV project for beginners to learn computer vision basics. In Single Object Tracking (SOT), the bounding box of the target in the first frame is given to the tracker. The goal of the tracker is then to locate the same target in all the other frames. If you need resources then In this 1-hour long project-based course, you will learn how to do Computer Vision on images with OpenCV and Python using Jupyter Notebook. This course runs on Coursera's hands-on project platform called Rhyme. The best thing about this project-based course is that you don't need to set up your development environment. For this project, you'll get instant access to a cloud desktop with Python, Jupyter, and OpenCV pre-installed.

66. OpenCV Project to Master Advanced Computer Vision Concepts

OpenCV is a huge open-source library for computer vision, machine learning, and image processing. It can process images and videos to identify objects, faces, or even the handwriting of a human. When integrated with various libraries, such as "NumPy," a highly optimized library for numerical operations, the number of weapons increases in your Arsenal, i.e., whatever operations one can do in NumPy can be combined with OpenCV.

67. Pencil Sketch with Python

In Python, an image is just a two-dimensional array of integers. So one can do a couple of matrix manipulations using various python modules to get some very interesting effects. To convert the normal image to a sketch, we will change its original RGB values and assign its RGB values similar to grey, in this way a sketch of the input image will be generated.

68. Plant Species Identification

Image-based methods are considered a promising approach for species identification. A user can take a picture of a plant in the field with the built-in camera of a mobile device and analyze it with an installed recognition application to identify the species or at least receive a list of possible species if a single match is impossible. By using a computer-aided plant identification system also non-professionals can take part in this process. Therefore, it is not surprising that large numbers of research studies are devoted to automating the plant species identification process.

69. Pneumonia Detection

With the growing technological advancements we have in this day and age, it is possible to use tools based on deep learning frameworks to detect pneumonia based on chest x-ray images. The challenge here would be to aid the diagnosis process which allows for expedited treatment and better clinical outcomes. The dataset that will be used for this project will be the Chest X-Ray Images (Pneumonia) from Kaggle. The dataset consists of training data, validation data, and testing data. The training data consists of 5,216 chest x-ray images with 3,875 images shown to have pneumonia and 1,341 images shown to be normal.

70. Predict Credit Default -Credit Risk Prediction Project

Credit default risk is simply known as the possibility of a loss for a lender due to a borrower's failure to repay a loan. Credit analysts are typically responsible for assessing this risk by thoroughly analyzing a borrower's capability to repay a loan — but long gone are the days of credit analysts, it's the machine learning age! Machine learning algorithms have a lot to offer to the world of credit risk assessment due to their unparalleled predictive power and speed. In this article, we will be utilizing machine learning's power to predict whether a borrower will default on a loan or not and to predict their probability of default.

71. Predict Tinder Matches with Machine Learning

The simple Tinder algorithm can swipe left and right based on the recommendations of a pre-trained deep neural network (Machine Learning). Convolutional Neural Networks are used in this process as they recognize objects, places, and people in your photos, signs, people, and lights in self-driving cars, crops, forests, and traffic in aerial imagery, various anomalies in medical images, and all kinds of other useful things. But once in a while, these powerful visual recognition models can also be warped for distraction, fun, and amusement.

72. Predict US Elections with Python

Two different models of two correlated datasets can be used for this project, one being a Multi Linear Regression model; the other a Classification model. Using the first model will give some sort of insight and direction as to what the result is supposed to look like. For the regression model, two dependent variables namely 'Democrats (year)' and 'Republicans (year)' can be used to test the model's accuracy.

73. Predicting Interest Levels of Rental Listings

Finding the perfect place to call your new home should be more than browsing through endless listings. RentHop makes apartment search smarter by using data to sort rental listings by quality. But while looking for the perfect apartment is difficult enough, structuring and making sense of all available real estate data programmatically is even harder. Two Sigma invites you to apply your talents in this recruiting competition featuring rental listing data from RentHop. We will predict the number of inquiries a new listing receives based on the listing's creation date and other features. Doing so will help RentHop better handle fraud control, identify potential listing quality issues, and allow owners and agents to better understand renters' needs and preferences.

74. Predicting Wine Quality using Wine Quality Dataset

The quality of wine can be predicted by using the wine quality dataset from Kaggle. This dataset has the fundamental features which are responsible for affecting the quality of the wine. By the use of several Machine learning models, we will predict the quality of the wine. Here we will

only deal with the white type wine quality, we use classification techniques to check further the quality of the wine i.e. is it good or bad.

75. Production Line Performance Checker

Bosch, one of the world's leading manufacturing companies, has an imperative to ensure that the recipes for the production of its advanced mechanical components are

of the highest quality and safety standards. Part of doing so is closely monitoring its parts as they progress through the manufacturing processes. Because Bosch records data at every step along its assembly lines, it can apply advanced analytics to improve these manufacturing processes. However, the intricacies of the data and the complexities of the production line pose problems for current methods. In this competition, Bosch is challenging Kagglers to predict internal failures using thousands of measurements and tests made for each component along the assembly line. This would enable Bosch to bring quality products at lower costs to the end-user.

76. Profit Prediction Model

The Predictive Revenue Model is a data analytic model which uses financial, marketing, and advertising data to predict the return on advertising within multiple scenarios. Capturing the initial marketing spend, the model represents the revenue flow from customer interaction at the top of the funnel.

77. Projection of Store Sales

Sales forecasting is the process of estimating future sales with the goal of better informing your decisions. A forecast is typically based on any combination of past sales data, industry benchmarks, or economic trends. It's a method designed to help you better manage your workforce, ash flow, and any other resources that may affect revenue and sales. It's typically easier for established businesses to create more accurate sales forecasts based on previous sales data. Newer businesses, on the other hand, will have to rely on market research, competitive benchmarks, and other forms of interest to establish a baseline for sales numbers.

78. Real-Time Face Mask Detection

Computer vision and image processing have an extraordinary impact on the detection of the face mask. Face detection has a range of case applications, from face recognition to facial movements, where the latter is required to show the face with extremely high accuracy. As machine learning algorithms progress rapidly, the threats posed by face mask detection technology still seem effectively handled. This innovation is becoming increasingly important as it is used to recognize faces in images and real-time video feeds. However, for the currently proposed models of face mask detection, face detection alone is a very tough task. In creating more improved facial detectors, following the remarkable results of current face detectors, the analysis of events and video surveillance is always challenging.

79. Real-time Gender Detection System

Gender detection is one of the popular computer vision applications. When you use a camera to detect a person's gender instead of detecting it on a picture, it can be said to be a real-time gender detection system. There are many libraries and frameworks in Python that can be used to create a real-time gender detection system. Some of these libraries include Yolo, Tensorflow, OpenCV, and Cvlib.

80. Real-time Sentiment Analysis

As the availability of the internet became easier, people started using social media platforms as the primary medium for sharing their opinions. Every day, millions of

opinions from different parts of the world are posted on Twitter. The primary goal of Twitter is to let people share their opinion with a big audience. So, if the authors can effectively analyze the tweets, valuable information can be gained. Storing these opinions in a structured manner and then using that to analyze people's reactions and perceptions about buying a product or a service is a very vital step for any corporate firm. Sentiment analysis aims to analyze and discover the sentiments behind opinions of various people on different subjects like commercial products, politics, and daily societal issues.

81. Recommender System Machine Learning Project for Beginners

A recommendation system is a type of algorithm designed to recommend or suggest things to the user based on many different factors. The recommendation system deals with a large amount of data and filters it out based on users' preferences and interests. With the rise of Youtube, Netflix, Amazon, etc., recommendation systems have taken a crucial place. Recommender systems are critical in many industries as they can help to generate a large amount of revenue. This project mainly focuses on the basics of the recommendation system and a brief introduction to the different algorithms. Implementation of a Rule-based recommendation system has also been covered.

82. Resume Parser NLP Spacy Python

This project uses Python's library, SpaCy to implement various NLP (natural language processing) techniques like tokenization, lemmatization, parts of speech tagging, etc., for building a resume parser in Python. And, considering all the resumes are submitted in PDF format, you will learn how to implement optical character recognition (OCR) for extracting textual data from the documents. The resulting application will require minimum human intervention to extract crucial information from a resume, such as an applicant's work experience, name, geographical location, etc.

83. Sales Prediction

Sales forecasting is a process of estimating future sales. Several companies conduct sales forecasts to enable their customers to make informed business decisions and predict short-term and long-term performance. This project would need its users to deploy Apache Spark Machine Learning using the Databricks platform community edition server which will allow the professionals to execute their spark code, free of cost on their servers just by registering through email id. Besides, there are several resources on the internet that might help professionals create this project accurately.

84. Sarcasm Detection

Sarcasm has been a part of our language since the beginning of time. Sarcasm detection is a natural language processing and binary classification task. The candidates executing the project can train a machine learning model to scan whether or not a sentence is sarcastic using a dataset of sarcastic and non-sarcastic sentences. Sarcasm detection using machine learning can be executed efficiently using Python programming language. The sarcasm detection project would include a dataset that would contain labels, used to predict sarcasm in a sentence.

85. Skin Cancer Classification

This project is about the detection and classification of various types of skin cancer using machine learning and image processing tools. Skin cancer is a deadly disease and early detection increases the survival rate. This project is aimed to build deep learning methods to classify dermal cell images and detect skin cancer. The results would yield almost accurate results that use deep learning algorithms in its core implementations, which are also used to construct models that assist in predicting abnormal skin cancer with improved accuracy.

86. Social Media Ads Classification

Classifying social media ads means analyzing their social media ads to find the most profitable customers for the products who are most likely to buy the product. Since not all products suit the target audience, using this project can help the enterprise leaders to determine whether a person will buy their product or not by classifying their social media ads. The dataset that will be used for the social media ads classification would contain information about a product's media advertising campaign that would enable the analysts to predict whether the target audience has purchased the product or not.

87. Social Media Followers Prediction

The activity of social media followers' prediction is quite crucial for content creators or business leaders. To predict the increase or decrease in the number of social media followers, professionals need to use a dataset that will yield the type of activities that the people on your social media perform. It is very difficult to find such a dataset on social media platforms like Facebook and Instagram since these platforms do not provide much data related to the followers. So, professionals need to find different sources from which they can gather this distinct information.

88. Social Media Sentiment Analysis

Sentiment analysis will help enterprise leaders perform a detailed opinion analysis that will help them understand public opinion on the company and its products. To gather such analytic information, you can start collecting data from the relevant sources like Facebook and Twitter and analyze the conversations between the users cand find the overall brand perception in the market. Using libraries offered by Python would be extremely helpful for those trying to analyze the sentiments of their followers. Scraping libraries like 'Beautiful Soup' will be fitting to collect data from websites.

89. Speech Emotion Recognition

Emotion recognition is quite a challenging task, but emotions are not constant. A speech emotion recognition system is a collection of methodologies that process and classify speech signals to detect emotions embedded in them. These SER systems can be used in a wide variety of application areas like an interactive voice-based assistant or caller-agent conversation analysis. The problem of speech emotion recognition can be solved by analyzing one or more of these features. With the help of this project, professionals can leverage machine learning to obtain the underlying emotion from speech audio data and some insights into the human expression of emotion through voice.

90. Stock Prices Predictor

Stock price prediction using machine learning is the process of predicting the future value of a stock traded on a stock exchange for reaping extreme profits. With the several multiple factors that are involved in predicting stock prices with high accuracy, machine learning plays a vital role in this. Professionals can use treating stock data using time-series data. The idea will be to measure the importance of recent and old data and determine which parameters are affecting the current prices of the stock or will affect it in the future.

91. Survival Prediction on the Titanic Ship

Survival prediction of the Titanic ship is a whole process of building a machine learning model on the Titanic dataset that is used by several people all over the world. It provides information on the fate of passengers on the Titanic, and the collection of data is summarized according to economic status, gender, age, and survival. This machine learning project has revealed that almost 60% of the passenger in the first class survived. Less than 30% of the passengers in the third class survived, which indicates that almost half of the passengers that were originally on the Titanic would have survived.

92. Ted Talks Recommendation System

Ted Talks is a good source to learn and take inspiration from. A Ted Talks Recommendation System has been entirely based on the content rather than based on the data of a user. As a user generally watches videos on Youtube and other applications mostly to get entertained, a Ted user generally watches the Ted Talks mainly to take inspiration, hence, user data is of no use here. To create such a system, researchers generally use the concept of cosine similarity in machine learning. With the integration of Python, you can definitely create this Ted Talks Recommendation System.

93. Topic Modeling with Python

Topic modeling lets developers implement helpful features like detecting breaking news on social media, recommending personalized messages, detecting fake users, and characterizing information flow. Text mining or topic modeling generally indicates the process of automatically identifying topics present in a text object and deriving hidden patterns exhibited by a text corpus. Hence, it better assists better decision-making. It is basically an unsupervised approach used for finding and observing a bunch of words in large clusters of texts.

94. Uber Data Analysis Project

Uber is responsible for moving data for people and making hassle-free deliveries. Uber collects large volumes of data and its fantastic team handles Uber data analysis using machine learning tools and frameworks. Creating this project would help Uber customers and riders to get positive experiences. But before starting with any machine learning project, one should know that it is essential to realize how prevalent the exercise of exploratory data analysis is in these projects.

95. Ultrasound Nerve Segmentation

Even the bravest people can get squeamish at the very name of surgeries. It can cause discomfort and surgical pain. But using this machine learning, cum data science project

one can identify nerve structures in a data set of ultrasound images. This will help enhance catheter placement and contribute to a more pain-free future for treatments. The creators of this project are working to improve pain management through the use of indwelling catheters that block or mitigate pain at the source.

96. Visualize a Solar System with Python

To complete this project, the participants can use the great selection of graphics libraries in Python which will allow them to plot all major types of graphics so that they can visualize all the data. Participants can use a combination of Plotly Scatter3D and Plotly Surface plots to create an entire solar system inside a graph. By creating more code lines using Python, this project can be recreated according to the desired results.

97. Whatsapp Group Chat Analysis

Whatsapp group chat analysis is a machine learning and data science project. The participants will receive a chat at the end of the project will not require any cleaning and preparation and can be used directly for the task. But before starting with the project, they will have to make sure that the data which is ready to use should have the format of the data and time of the messages changed, and it can be done extremely easily.

98. Worldwide Billionaires Analysis

The number of billionaires in a country speaks a lot about the business environment, startup success rate, and several other economic features of that country. So, if you are eager to find relationships among billionaires around the world, this project might be the perfect opportunity for you. With the help of this project, the participants can find patterns among billionaires around the world to analyze the business environment of countries. This would eventually help the success of a business or startup, depending on the accurate prediction of the business environment of the country.

99. Xbox Game Prediction Project

The data generated by people while surfing the internet can be used to predict the interest of the users. The BestBuy consumer electronics company has provided the data of millions of searches from users and will allow the participants to predict the Xbox game that a user will be most interested to buy.

100. Zillow Home Value Prediction ML Project

Zillow is a popular online real estate and rental marketplace dedicated to providing customers with data to make the best possible housing decision. The Zestimate was initially created to give consumers as much information as possible about homes and the housing market, marking the first time consumers had access to the type of home value information at no cost. This project would essentially demonstrate to improve the Zestimate residual error which is also called 'log error' that would in turn provide the actual sales price.

A) COURSE-BASED PROJECT (Psychomotor skills)

To enhance their skill set in the integrated course, the students are advised to execute course-

SNo	Suggested Projects	BTL
1	Worldwide Billionaires Analysis	K3
2	Unemployment Analysis	K3
3	Car Price Prediction Model	K3
4	Spam Detection	K3
5	Count Objects in Image	K3
6	WhatsApp Chats Sentiment Analysis	K3
7	WhatsApp Chats Analysis	K3
8	Microsoft Stock Price Prediction	K3
9	Covid-19 Vaccine Analysis	K3
10	Video Game Sales Prediction Model	K3
11	Student Grades Prediction Model	K3
12	Saving a Machine Learning Model	K3
13	Uber Trips Analysis	K3
14	Google Search Analysis	K3
15	Tesla Stock Price Prediction Model	K3
16	Financial Budget Analysis	K3
17	Click-Through Rate Prediction Model	K3
18	Interactive Language Translator	K3
19	Language Detection	K3
20	Create a Chatbot with Python	K3
21	Best Streaming Service Analysis	K3
22	Stock Price Prediction	K3
23	Data Science Project on President Heights	K3
24	Data Science Project on Birth Rate Analysis	K3
25	Data Science Project on Time Series	K3
26	Data Science Project on Area and Population	K3
27	A Complete Machine Learning Project Walkthrough	K3
28	Text Summarization	K3
29	Keyword Extraction	K3
30	Data Science Projects on Finance	K3
31	Data Science Projects on Marketing	K3
32	End to end Machine Learning Projects	K3
33	Data Science Projects -Advanced	K3
34	Amazon Product Reviews Sentiment Analysis	K3
35	Hate Speech Detection	K3
36	End-to-end Hate Speech Detection System	K3
37	End-to-end Fake News Detection System	K3
38	End-to-end Spam Detection System	K3
39	Hotel Reviews Sentiment Analysis	K3
40	Real-time Gender Detection System	K3

41	Dogecoin Price Prediction	K3
42	Google Play Store Sentiment Analysis	K3
43	Amazon Alexa Reviews Sentiment Analysis	K3
44	Social Media Ads Classification	K3
45	Fake News Detection	K3
46	End-to-End Machine Learning Model	K3
47	Gender Detection	K3
48	Sales Prediction	K3
49	Currency Exchange Rate Prediction	K3
50	End-to-end Machine Learning Project	K3
51	Profit Prediction Model	K3
52	Automatic Time Series Forecasting	K3
53	Ted-Talks Recommendation System	K3
54	Real-time Sentiment Analysis	K3
55	Amazon Recommendation System	K3
56	Mobile Price Classification	K3
57	Spotify Recommendation System	K3
58	Text Emotions Detection	K3
59	Hotel Recommendation System	K3
60	Bankruptcy Prediction Model	K3
61	Customer Personality Analysis	K3
62	Instagram Algorithm	K3
63	Diamond Price Prediction Model	K3
64	Book Recommendation System	K3
65	Netflix Data Analysis	K3
66	Satellite Imagery Analysis	K3
67	Topic Modelling	K3
68	Covid-19 Vaccine Sentiment Analysis	K3
69	Human Activity Recognition Model	K3
70	Cohort Analysis	K3
71	End to end Machine Learning Project	K3
72	Life Expectancy Analysis	K3
73	Bitcoin Price Prediction	K3
74	Human Resource Analysis	K3
75	Energy Consumption Prediction	K3
76	LightGBM Model	K3
77	FastAI Model	K3
78	House Price Prediction	K3
79	Real-Time Face Mask Detection	K3
80	Netflix Recommendation System	K3
81	Named Entity Recognition	K3
82	Number Plate Detection	K3

83	IPL Analysis with Python	K3
84	Gold Price Prediction	K3
85	Object Detection	K3
86	Highest-Paid Athletes Analysis	K3
87	Text Generation	K3
88	Spelling Correction with Python	K3
89	Income Classification	K3
90	Customer Churn Prediction	K3
91	Language Translation Model	K3
92	Resume Screening	K3
93	Sign Language Classification	K3
94	Online Shopping Intention Analysis	K3
95	Network Graph Analysis	K3
96	Keyword Extraction	K3
97	Amazon Best Selling Books Analysis	K3
98	Covid-19 Cases Prediction for Next 30 Days	K3
99	YouTube Trending Videos Analysis	K3
100	Gender Classification	K3
101	Flower Recognition	K3
102	Pneumonia Detection	K3
103	Employee Attrition Prediction	K3
104	Hand Gesture Recognition	K3
105	Face mask Detection	K3
106	Market Basket Analysis using Apriori Algorithm	K3
107	Breast Cancer Detection	K3
108	Earthquake Prediction Model	K3
109	Outlier Detection	K3
110	Heart Disease Prediction	K3
111	Plastic Users Analysis	K3
112	Landmark Detection	K3
113	Chatbot with Machine Learning	K3
114	Next Word Prediction Model	K3
115	Age and Gender Detection with Python	K3
116	Autocorrect Keyboard with Python and Machine Learning.	K3
117	Machine Learning in 5 lines of code.	K3
118	Deepfake Detection with Machine Learning.	K3
119	Predict US Elections with Python.	K3
120	Fake Currency Detection with Machine Learning.	K3

121	Predict Tinder Matches with Machine Learning.	K3
122	Image Segmentation	K3
123	Title Generator with Python.	K3
124	Build Google Search Algorithm with Python.	K3
125	Face Landmarks Detection	K3
126	Pencil Sketch with Python.	K3
127	OpenAI Gym in Machine Learning	K3
128	Machine Translation Model	K3
129	Covid 19 Analysis.	K3
130	Build the TikTok Algorithm with Machine Learning.	K3
131	Analyze ILP Matches.	K3
132	Barcode and QR code Reader with Python	K3
133	Extract Text From PDF with Python.	K3
134	Predict IPL Winner 2020.	K3
135	Predict Car Prices.	K3
136	Analyze Call Records.	K3
137	Create an API with Python.	K3
138	Send Custom Emails with Python.	K3
139	Colour Recognition with Machine Learning.	K3
140	Create a 3D Video Animation.	K3
141	Graph Algorithms in Machine Learning.	K3
142	Image Features Extraction.	K3
143	Rainfall Prediction with Machine Learning.	K3
144	Classify Nationalities with Machine Learning.	K3
145	Fraud Detection with Machine Learning.	K3
146	Topic Modeling with Python	K3
147	Named Entity Recognition	K3
148	WhatsApp Group Chat Analysis	K3
149	Translate Languages Using Python	K3
150	Network Security Analysis	K3
151	Bar Chart Race with Python	K3
152	Keyword Research with Python	K3
153	Fashion Recommendation System	K3
154	Face Detection with Python	K3
155	Emotion Detection Model	K3
156	Telegram Bot with Python	K3

157	Handwriting Recognition	K3
158	Visualize a Solar System with Python.	K3
159	Hate Speech Detection Model	K3
160	Build Instagram Filters with Python.	K3
161	Contact Tracing with Machine Learning	K3
162	Deploy a Chatbot with Python into a Web Application	K3
163	Summarize Text with Machine Learning	K3
164	Language Classification with Machine Learning.	K3
165	OTP Verification GUI with Python	K3
166	Create an Audiobook with Python	K3
167	Titanic Survival Analysis	K3
168	Image Recognition with PyTorch	K3
169	Gender Classification Model	K3
170	Daily Births Forecasting	K3
171	Binary Search Algorithm	K3
172	Galaxy Classification with Machine Learning	K3
173	Time Series with LSTM Model	K3
174	Image Classification with TensorFlow	K3
175	Predict Weather with Machine Learning	K3
176	Create a Package with Python	K3
177	Computer Vision with Python	K3
178	Diamonds Analysis with Python	K3
179	Text Classification with TensorFlow	K3
180	Image Segmentation with Python	K3
181	Ridge and Lasso Regression	K3
182	Predict Fuel Efficiency	K3
183	ABC Analysis with Python	K3
184	Lung Segmentation with Machine Learning	K3
185	RFM Analysis with Python	K3
186	Build Neural Network with Python Code	K3
187	Genetic Algorithm with Python	K3
188	Predict Migration with Machine Learning	K3
189	Analyze Healthcare Data	K3
190	NLP For Other Languages	K3
191	Data Science Project on Text and Annotations	K3
192	Data Science Project on — Handwritten Digits	K3

193	Data Science Project on- Support Vector Machines	K3
194	Data Science Project — Stock Price Prediction with Machine Learning	K3
195	Data Science Project on — Classification of Text	K3
196	Data Science Project on-Extracting HOG Features	K3
197	Data Science Project on — Amazon Products Reviews Sentiment Analysis with Machine Learning	K3
198	Data Science Project — Email spam Detection with Machine Learning	K3
199	Data Science Project — Heart Disease Prediction with Machine Learning	K3
200	Data Science Project — Movie Recommendation System	K3
201	Data Science Project — Student Performance Analysis with Machine Learning	K3
202	Data Science Project on FIFA Analysis with python	K3
203	Data science project — Fake News Classification	K3
204	Data Science Project — DNA Sequencing with Machine Learning	K3
205	Data Science Project — Book Recommendation System with Machine Learning	K3
206	Data Science Project — Bitcoin Price Prediction with Machine Learning	K3
207	Machine Learning Project — Artificial Neural Networks	K3
208	Algorithmic Trading Strategy with Machine Learning and Python	K3
209	Movie Reviews Sentiment Analysis -Binary Classification with Machine Learning	K3
210	Data Science Project — Supermarket Sales Analysis	K3
211	Data Science Project — GDP Analysis	K3
212	Data Science Project — Predict Customer Churn with Python and Machine Learning	K3

213	Data Science Project — San Francisco Crime Analysis	K3
214	Machine Learning Project — Human Activity Recognition using Smartphone Data	K3
215	Credit Card Fraud Detection with Machine Learning	K3
216	Weather Forecasting with Machine Learning	K3
217	SMS Spam Detection with Machine Learning	K3
218	Covid-19 Detection with Machine Learning and AI	K3
219	Customer Segmentation with Machine Learning	K3
220	Employee Turnover Prediction with Machine Learning	K3
221	Predict Diabetes with Machine Learning	K3
222	Image Classification with PyTorch	K3
223	Time Series Forecasting with ARIMA Model	K3
224	Natural Language Processing on WhatsApp Chats	K3
225	Fake News Detection Model	K3
226	Image Classification with Artificial Neural Networks	K3
227	Binary Classification Model	K3
228	Data Augmentation with Deep Learning	K3
229	Next Word Prediction Model.	K3
230	Image Segmentation	K3
231	XGBoost Algorithm in Machine Learning	K3
232	Face Landmarks Detection	K3
233	Image Filtering with Machine Learning	K3
234	Audio Feature Extraction	K3
235	Machine Translation Model	K3
236	Gender Classification Model	K3
237	Create a 3D Video with Python and Machine Learning.	K3
238	Named Entity Recognition	K3
239	WhatsApp Group Chat Analysis	K3
240	Translate Languages Using Python	K3
241	Covid-19 Projects with Machine Learning	K3

242	Deep Learning Projects with Machine Learning	K3
243	Recommendation System Projects with Machine Learning	K3
244	Sentiment Analysis Projects with Machine Learning	K3
245	Classification Projects For Machine Learning	K3
246	Regression Projects for Machine Learning	K3
247	Chatbot Projects with Python	K3
248	Machine Learning Projects on Future Prediction	K3
249	Machine Learning Projects for Resume	K3
250	Best Data Science Projects for Resume	K3
251	Machine Learning Projects on Social Media Analysis	K3

based
design projects. Some sample projects are given below:

B) SELF-LEARNING THROUGH MOOCs (Cognitive Skills): Certification
1. https://www.udemy.com/topic/machine-learning/
2. https://www.coursera.org/learn/machine-learning
3. https://developers.google.com/machine-learning/crash-course
4. https://www.edx.org/learn/machine-learning