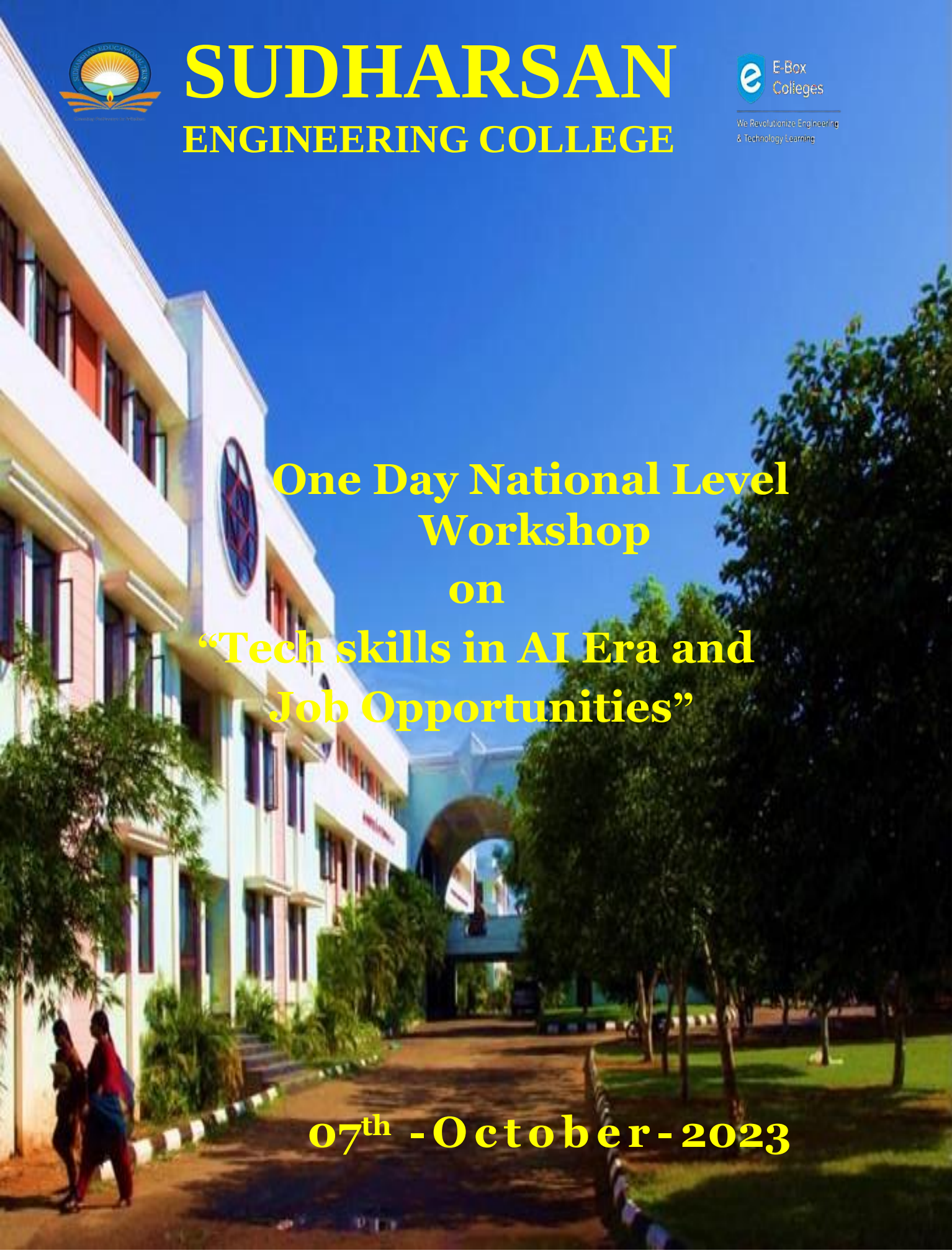
SUDHARSAN
ENGINEERING COLLEGE

E-Box Colleges
We Revolutionize Engineering & Technology Learning



One Day National Level Workshop on "Tech skills in AI Era and Job Opportunities"

07th -October-2023

SUDHARSAN ENGINEERING COLLEGE

Sathiyamangalam, Pudukkottai.

One Day National Level Workshop on

"Tech Skills in AI Era and Job Opportunities"

Organized By

Department of CSE & AIDS

Date:7th OCTOBER 2023

Report Prepared By

1. **Mrs.P.Parvathi ASP/CSE**
2. **Mrs.R.Kavitha AP/AI&DS**

SUDHARSAN ENGINEERING COLLEGE

Sathiyamangalam, Pudukkottai.

One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

POSTER

SUDHARSAN ENGINEERING COLLEGE

Sathiyamangalam, Pudukkottai.

E-Box Colleges

One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

WORKSHOP SCHEDULE

TIME	TOPIC	SPEAKER
10.00 am TO 10.10 am	**Welcome Address**	**Mr.S.John Joseph HOD – AI&DS**
10.10 AM TO 10.20 AM	**Chief Guest Introduction**	**Mrs.P.Parvathi ASP / CSE**
10.20 AM TO 11.00 AM	***Topics to be covered*** ***1. Data Science / Data Engineering*** • Types / Categories of Data • Data Analysis &Data Mining • Predictive Modeling ***2. Machine Learning (ML)*** • Supervised & Unsupervised Learning • Regression & Classification • Clustering • Dimensionality Reduction • Ensemble Learning • Model Evaluation	**Dr.R.Amarnath Senior Architect HCL Technologies**
11.00 am – 11.15 am	**BREAK**	

TIME	TOPIC	SPEAKER
11.15 AM – 1.00 PM	***3. Deep Learning (DL)*** • Neural Networks • CNNs (Convolutional Neural Networks) • RNNs (Recurrent Neural Networks) • NLP (Natural Language Processing) **4. Introduction** – LLM (Large LanguageModel) - Like ChatGPT, Bard, llama, etc **5. Research & Applications inVarious Domains** EdTech, FinTech, Agriculture,Medical, Automobiles	
1.00 PM to 2.00 PM	**LUNCH**	
2.00PM to 4.30 PM	**Hands on Training**	**Dr.T.Punitha Director Amphi Ventures**
4.30 PM	**Vote of Thanks**	**Mrs.R.Kavitha AP/AI&DS**

One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

EXTERNAL PARTICIPANTS LIST

S.NO	NAME	FACULTY / STUDENT	COLLEGE NAME
1	G. VISHNU RITHAN	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
2	N.ARAVINTH	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
3	MRS.R.PREETHI	ASSISTANT PROFESSOR	SHANMUGANATHAN ENGINEERING COLLEGE
4	MR. T. MARUDHAN	ASSOCIATE PROFESSOR	ALAGAPPA CHETTIAR GOVERNMENT COLLEGE OF ENGINEERING AND TECHNOLOGY
5	PROF. R. THIRUMAVALAVAN	ASSOCIATE PROFESSOR	MOUNT ZION COLLEGE OF ENGINEERING AND TECHNOLOGY
6	SASIREKA G	ASSISTANT PROFESSOR	SHANMUGANATHAN ENGINEERING COLLEGE
7	MRS.P.SRILEGA	ASSISTANT PROFESSOR	EXCEL ENGINEERING COLLEGE
8	MRS C.PRATHIPA	ASSISTANT PROFESSOR	EXCEL ENGINEERING COLLEGE
9	P.DHARSHINI	STUDENT	UNIVERSITY COLLEGE OF ENGINEERING, ANNA UNIVERSITY, BIT CAMPUS,TRICHY
10	MS.A.AKSHAYA.	STUDENT	SIR ISSAC NEWTON COLLEGE OF ENGINEERING AND TECHNOLOGY
11	VIDHYA	ASSISTANT PROFESSOR	INDIRA GANESAN COLLEGE OF ENGINEERING
12	MRS.S. LAVANYA PRABHA	ASSISTANT PROFESSOR	SRI BHARATHI ENGINEERING COLLEGE FOR WOMEN
13	S. SHIYAMALADEVI	STUDENT	SIR ISSAC NEWTON COLLEGE OF ENGINEERING AND TECHNOLOGY
14	R. BHARATH KUMAR	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
15	M.GANGA LAKSHMI	STUDENT	CHENDHURAN ENGINEERING COLLEGE
16	MRS.V.RAMYA	ASSISTANT PROFESSOR	EXCEL ENGINEERING COLLEGE
17	J.OBURADHA	ASSISTANT PROFESSOR	EXCEL ENGINEERING COLLEGE
18	K.PREMA	ASSISTANT PROFESSOR	EXCEL ENGINEERING COLLEGE
19	MRS.M.VIDYA	ASSISTANT PROFESSOR	EXCEL ENGINEERING COLLEGE
20	MRS.M.SINDHU	ASSISTANT PROFESSOR	SRI RAAJA RAAJAN COLLEGE OF ENGINEERING AND TECHNOLOGY

EXTERNAL PARTICIPANTS LIST

S.NO	NAME	FACULTY / STUDENT	COLLEGE NAME
21	MS. S. SRINIDHI	STUDENT	CHENDHURAN COLLEGE OF ENGINEERING AND TECHNOLOGY
22	C. SHALINI	STUDENT	CHETTINAD COLLEGE OF ENGINEERING AND TECHNOLOGY
23	N.NAGARAJAN	STUDENT	MAHATH AMMA INSTITUTE OF TECHNOLOGY AND ENGINEERING
24	R.SANGEETHA	STUDENT	MOTHER TERASA COLLEGE OF ENGINEERING AND TECHNOLOGY
25	VAITHEESWARAN CHANDRASEKAR	STUDENT	ARUMUGAM PILLAI SEETHAI AMMAL COLLEGE
26	MRS.S. HIMAGIRI SUDHA	ASSISTANT PROFESSOR	MOOKAMBIKAI COLLEGE OF ENGINEERING
27	MRS. B. JULIET CELINE MARY	ASSISTANT PROFESSOR	MOOKAMBIGAI COLLEGE OF ENGINEERING
28	MRS. C. GOHILA PRIYADHARSHINI	ASSISTANT PROFESSOR	MOOKAMBIGAI COLLEGE OF ENGINEERING
29	B.JAVAHAR NISHA	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
30	MS.R.DHARANI	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
31	B. DIVYA	ASSISTANT PROFESSOR	INDRA GANESAN COLLEGE OF ENGINEERIMG
32	MR. M. VIJAY	RESEARCH SCHOLAR	THIAGARAJAR COLLEGE OF ENGINEERING
33	R.MURUGESAN	RESEARCH SCHOLAR	PAAVAI COLLEGE OF ENGINEERING
34	M.KALAISELVI	ASSISTANT PROFESSOR	CARE COLLEGE OF ENGINEERING
35	RITHIK ROSHAN B	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
36	M .VISVESHVARAN	STUDENT	SRM TRP ENGINEERING COLLEGE
37	MR.S.THAMEEM ANSARI	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
38	THRISHA K	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
39	SAJITHA PARVEEN.P	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
40	MR.B.BRAMEASWARAN	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE

EXTERNAL PARTICIPANTS LIST

S.NO	NAME	FACULTY / STUDENT	COLLEGE NAME
41	RAKESH S	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
42	MR. U. KISHORE KUMAR	STUDENT	M.A.M. COLLEGE OF ENGINEERING
43	M.VIJAYALAKSHMI	ASSISTANT PROFESSOR	CARE COLLEGE OF ENGINEERING
44	ARUNKUMAR M	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
45	MOHAMED FAHADH A	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
46	AAKASH.S	STUDENT	MOOKAMBIGAI COLLEGE OF ENGINEERING
47	S. KAYATHRI	STUDENT	INDRA GANESAN COLLEGE OF ENGINEERING
48	SANTHOSH. S	STUDENT	MAHATH AMMA INSTITUTE OF ENGINEERING AND TECHNOLOGY
49	MR. M. DINESH KUMAR	STUDENT	CHENDHURAN POLYTECHNIC COLLEGE
50	MR.V. DHINESH KUMAR	STUDENT	CHENDHURAN POLYTECHNIC COLLELLE
51	MR.A.MOHAMMED ARABAATH ALI	STUDENT	CHEDHURAN POLYTEHNIC COLLEGE
52	M. RAJALAKSHMI	ASSISTANT PROFESSOR	DR. MAHALINGAM COLLEGE OF ENGINEERING AND TECHNOLOGY
53	K. MEGALA	STUDENT	IFET COLLEGE OF ENGINEERING
54	S.KAVIHARSHINI	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
55	DEIVANAI R	STUDENT	IFET COLLEGE OF ENGINEERING
56	M.ABIRAMI	STUDENT	SHANMUGANATHAN ENGINEERING COLLEGE
57	V. THIRUVIGNESH	STUDENT	URUMU DHANALAKSHMI COLLEGE
58	N.LAKKAIAN	STUDENT	MOTHER TERESA COLLEGE OF ENGINEERING AND TECHNOLOGY
59	N.MAHENDRAN	STUDENT	GOVERNMENT ARTS COLLEGE
60	K.VELMANI	STUDENT	GOVERNMENT ARTS COLLEGE,TRICHY-22
61	R.DHAVAMANI	STUDENT	GOVERNMENT ARTS COLLEGE
62	C.SANTHOSH	STUDENT	URUMUDHANALAKSHMI COLLEGE

One Day National Level Workshop
on
"Tech Skills in AI Era and Job Opportunities"

Online Google Meet Link

Dear Sir / Madam,

Greetings from Sudharsan Engineering College -Pudukkottai !!!

We appreciate your interest in registering the **"One day national level virtual workshop on Tech skills in AI Era and Job Opportunities 07.10.2023 (Saturday)".** In this regard, we expect your fullest support and co-operation in attending the Session.

Google meet Link for the Session - https://meet.google.com/qpo-pbno-yss

Date - 07-10-2023 (Saturday)
Time - 09:30 AM - 4:30 PM

Contact No:
+9894807239-Mrs.R.KAVITHA AP/AIDS
+9786367841-Mrs.P.PARVATHI ASP/CSE

One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

Report

Session 1: October 07, 2023

Presenter: Dr. R.Amarnath
Senior Architect
HCL Technologies

1. Data Science / Data Engineering

- **Types / Categories of Data:** Data Types are an important concept of statistics, which needs to be understood, to correctly apply statistical measurements to your data and therefore to correctly conclude certain assumptions about it. This blog post will introduce you to the different data types you need to know, to do proper exploratory data analysis (EDA), which is one of the most underestimated parts of a machine learning project.

- **Data Analysis :**Data analysis is a process for obtaining raw data, and subsequently converting it into information useful for decision-making by users. Data is collected and analyzed to answer questions, test hypotheses, or disprove theories.

- **Data Mining:**Data mining is the process of searching and analyzing a large batch of raw data in order to identify patterns and extract useful information. Companies use data mining software to learn more about their customers. It can help them to develop more effective marketing strategies, increase sales, and decrease costs.

- **Predictive Modeling:**Predictive modeling is a mathematical process used to predict future events or outcomes by analyzing patterns in a given set of input data. It is a crucial component of predictive analytics, a type of data analytics which uses current and historical data to forecast activity, behavior and trends.

2. Machine Learning (ML)

- **Supervised Learning :** Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence. It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately.

- **Unsupervised Learning:** Unsupervised learning, also known as unsupervised machine learning, uses machine learning algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention.

- **Semi-Supervised Learning:** Semi-supervised learning is a type of machine learning that falls in between supervised and unsupervised learning. It is a method that uses a small amount of labeled data and a large amount of unlabeled data to train a model.

- **Reinforcement Learning:** Reinforcement learning (RL) is an area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward.

- **Regression:** Machine Learning Regression is a technique for investigating the relationship between independent variables or features and a dependent variable or outcome. It's used as a method for predictive modelling in machine learning, in which an algorithm is used to predict continuous outcomes.

- **Classification :** Classification is a supervised machine learning method where the model tries to predict the correct label of a given input data. In classification, the model is fully trained using the training data, and then it is evaluated on test data before being used to perform prediction on new unseen data.

- **Clustering :** In machine learning too, we often group examples as a first step to understand a subject (data set) in a machine learning system. Grouping unlabeled examples is called clustering. As the examples are unlabeled, clustering relies on unsupervised machine learning.

- **Dimensionality Reduction:** Dimensionality reduction is the process of reducing the number of features in a dataset while retaining as much information as possible. This can be done to reduce the complexity of a model, improve the performance of a learning algorithm, or make it easier to visualize the data.

- **Ensemble Learning :** Ensemble learning helps improve machine learning results by combining several models. This approach allows the production of better predictive performance compared to a single model. Basic idea is to learn a set of classifiers (experts) and to allow them to vote. Advantage : Improvement in predictive accuracy.

- **Model Evaluation:** Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance, as well as its strengths and weaknesses. Model evaluation is important to assess the efficacy of a model during initial research phases, and it also plays a role in model monitoring.

3. Deep Learning (DL)

- **Neural Networks:** A neural network is a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. It is a type of machine learning process, called deep learning, that uses interconnected nodes or neurons in a layered structure that resembles the human brain.

- **CNNs (Convolutional Neural Networks) :** A CNN is a kind of network architecture for deep learning algorithms and is specifically used for image recognition and tasks that involve the processing of pixel data. There are other types of neural networks in deep learning, but for identifying and recognizing objects, CNNs are the network architecture of choice.

- **RNNs (Recurrent Neural Networks):** A recurrent neural network (RNN) is a type of artificial neural network which uses sequential data or time series data.

- **NLP (Natural Language Processing):** Natural language processing (NLP) combines computational linguistics, machine learning, and deep learning models to process human language. Computational linguistics. Computational linguistics is the science of understanding and constructing human language models with computers and software tools.

4. Introduction

LLM (Large LanguageModel)

- ChatGPT: ChatGPT is an artificial intelligence (AI) chatbot that uses natural language processing to create humanlike conversational dialogue. The language model can respond to questions and compose various written content, including articles, social media posts, essays, code and emails.

- Bard: Google Bard is an artificial intelligence chatbot that can respond to a user's questions (or prompts) on any subject with an almost human-like "understanding." It performs text-based tasks like creating various forms of content, summarizing text, and translating between languages.

- Llama: Meta has introduced Code Llama, a large language model capable of generating code from text prompts. Code Llama includes three versions with different sizes and specialized capabilities. The model has demonstrated impressive benchmark performance, outpacing other publicly available code-specific LLMs

5. Research & Applications in Various Domains

- **EdTech:** EdTech apps, short for "education technology" applications, are software designed for the education industry. By utilizing technology, they can enhance the quality of teaching and learning. These innovative tools have become popular. They cater to the diverse needs of students, educators, and educational institutions.

- **FinTech:** Fintech software encompasses diverse applications like mobile banking platforms, crowdfunded investment tools, cryptocurrency trading applications or robo-advisors - making processes more user-friendly, efficient and swift for customers.

- **Agriculture:** Agricultural Applications means (i) copying of animals to produce embryos and live animals; (ii) the genetic modifications of animals to produce genetically altered embryos and live animals; (iii) the generation of breeding lines of animals; (iv) animal cell line inventories; and (v) the production of animals.

- **Medical:** An application domain is the segment of reality for which a software system is developed. It is the background or starting point for the actual-state analysis and the creation of a domain model. An application domain can be an organization, a department within an organization, or a single workplace.

- **Automobiles:** automobile, byname auto, also called motorcar or car, a usually four-wheeled vehicle designed primarily for passenger transportation and commonly propelled by an internal-combustion engine using a volatile fuel.

Key Takeaways:

- Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans

- Deep learning is a key technology behind driverless cars, enabling them to recognize a stop sign, or to distinguish a pedestrian from a lamppost.

- Understanding Machine Learning.

- Machine learning is a subset of artificial intelligence that involves training machines to learn from data, without being explicitly programmed. …

Conclusion: In conclusion, deep learning is a significant advancement in the field of artificial intelligence, with the ability to process large amounts of data, achieve high levels of accuracy,

and adapt and improve over time.

SUDHARSAN ENGINEERING COLLEGE

Sathiyamangalam, Pudukkottai.

One Day National Level Workshop
on
"Tech Skills in AI Era and Job Opportunities"

Report

Session 2 : Hands on Training

Topic: Machine Learning

Presenter: Dr.T.Punitha,
Director,
Amphi Ventures

1. Data Exploration and Visualization

Question 1: Data exploration and visualization

```
#Import the required libraries
import pandas as pd
pd.set_option('display.max_columns',100)
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.simplefilter('ignore')

#Load the dataset
data = pd.read_csv("vehicle_data_1.csv")

#Step 1 : Shape or Dimensions of the DataFrame
print("DF Shape")
print(data.shape)

#Step 2 : Explore Data Types of the various columns
print("DF Column Types")
print(data.dtypes)

#Step 3 : Explore data by displaying 5 rows
print("Top 5 rows")
print(data.head())

#Numerical Features Plots
#create df of numeric features only.
num_cols = data.select_dtypes(exclude='object')
```

```
#Set size of figure
cols = 3
rows = 3
fig = plt.figure(figsize=(cols*5, rows*5))
plt.subplots_adjust(hspace=0.5, wspace=0.5)

#Step 4 : Plot  Histogram for all numeric features (Use sns.histplot())
#(Set title as "Histogram" .Save in file uahistogram.png)
plt.suptitle("Histogram")
for i, col in enumerate(num_cols):
   ax=fig.add_subplot(rows,cols,i+1)
   sns.histplot(x = data[col], ax = ax)
plt.savefig("uahistogram.png")
plt.clf()

#Step 5 : Plot KDE plot for all numeric features (Use sns.kdeplot())
#(Set title as "KDE Plot" .Save in file uakdeplot.png)
plt.suptitle("KDE Plot")
for i, col in enumerate(num_cols):
        ax=fig.add_subplot(rows,cols,i+1)
        sns.kdeplot(x=data[col],ax=ax)
plt.savefig("uakdeplot.png")
plt.clf()

#Step 6 : Plot Rug Plot for all numeric features (Use sns.rugplot())
#(Set title as "Rug Plot" .Save in file uarugplot.png)
plt.suptitle("Rug Plot")
for i, col in enumerate(num_cols):
        ax=fig.add_subplot(rows,cols,i+1)
```

```
        sns.rugplot(x=data[col],ax=ax)
plt.savefig("uarugplot.png")
plt.clf()

#Step 7 : Plot Box Plot for all numeric features (Use sns.boxplot())
#(Set title as "Box Plot" .Save in file uaboxplot.png)
plt.suptitle("Box Plot")
for i, col in enumerate(num_cols):
        ax=fig.add_subplot(rows,cols,i+1)
        sns.boxplot(x=data[col],ax=ax)
plt.savefig("uaboxplot.png")
plt.clf()

#Step 8 : Plot Violin Plot for all numeric features (Use sns.violinplot())
#(Set title as "Violin Plot" .Save in file uaviolinplot.png)
plt.suptitle("Violin Plot")
for i, col in enumerate(num_cols):
        ax=fig.add_subplot(rows,cols,i+1)
        sns.violinplot(x=data[col],ax=ax)
plt.savefig("uaviolinplot.png")
plt.clf()

#Step 9 : Plot Strip Plot for all numeric features (Use sns.stripplot())
#(Set title as "Strip Plot" .Save in file uastripplot.png)
plt.suptitle("Strip Plot")
for i, col in enumerate(num_cols):
        ax=fig.add_subplot(rows,cols,i+1)
        sns.stripplot(x=data[col],ax=ax)
plt.savefig("uastripplot.png")
```

```
plt.clf()

#Categorical Features Plots
#create df of categoric features only.
all_cat_cols = data.select_dtypes(include='object')

#Create df of categorical features where the number of unique values is less than 10
cat_cols = all_cat_cols.columns[all_cat_cols.nunique()<10]

#Set size of figure
cols = 2
rows = 2
fig = plt.figure(figsize=(cols*5, rows*5))
plt.subplots_adjust(hspace=0.5,wspace=0.5)

#Step 10 : Plot Count Plot for all Categoric features (Use sns.countplot())
#(Set title as "Count Plot" .Save in file uacountplot.png)
plt.suptitle("Count Plot")
for i, col in enumerate(cat_cols):
        ax=fig.add_subplot(rows,cols,i+1)
        sns.countplot(x=data[col],ax=ax)
plt.savefig("uacountplot.png")
plt.clf()

#Step 11 : Plot Pie Charts for all Categoric features (Use plt.pie())
plt.suptitle("Pie Chart")
for i, col in enumerate(cat_cols):
        df = data[col].value_counts()
        ax = fig.add_subplot(rows, cols, i+1)
```

```
        plt.pie(df, labels=df.index, autopct="%.0f%%")
plt.savefig("uapiechart.png")
plt.clf()
```

2. Data Preprocessing

Question 2: Data preprocessing

```
import pandas as pd
pd.set_option('display.max_columns',100)
from sklearn.model_selection import train_test_split

#Load data in Pandas
df = pd.read_csv('titanic_data.csv')

#Print columns and their datatypes
print("Columns")
print(df.dtypes)

#Display first 5 rows of data
print("First 5 rows of data")
print(df.head())

#Drop Columns --- Name, Ticket and Cabin
df=df.drop(columns=['Name','Ticket','Cabin'],axis=1)

#Print columns and their datatypes
print("Columns after dropping 3 cols")
print(df.dtypes)

#Create Dummy Variables
#Convert the Pclass, Sex and Embarked to columns in Pandas and drop them after conversion.
cols=['Pclass','Sex','Embarked']
d=[]
for i in cols:
```

```
    d.append(pd.get_dummies(df[i]))
ddf=pd.concat(d,axis=1)
df=pd.concat((df,ddf),axis=1)
df=df.drop(columns=['Pclass','Sex',"Embarked"],axis=1)

#Print columns and their datatypes
print("Columns after creating dummy variables")
print(df.dtypes)

#Print first 10 rows of the data. Observe missing values in age column
print("First 10 rows of data")
print(df.head(10))

#Handle missing rows in age by using interpolate
df['Age'] =df["Age"].interpolate()

#Print first 10 rows of the data. Observe the age column again
print("First 10 rows of data")
print(df.head(10))

#Convert DataFrame from Pandas to Numpy and separate out input features(X) and the output(y)
y=df['Survived'].values
x=(df.drop('Survived',axis=1)).values
#Split the dataset into training data and test data using train_test_split method
#Parameters used : test_size=0.3, shuffle=False, random_state=0
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.3,shuffle=False, random_state=0)

#Print First 5 rows of Training Set Features
print("First 5 rows of Training Set Features")
```

```
print(x_train[:5,:])


#Print First 5 rows of Testing Set Features
print("First 5 rows of Testing Set Features")
print(x_test[:5,:])
```

3. Machine Learning Classification

Question 3: ML / Classification

```
#Step 1: Import the required libraries
import numpy as np
import pandas as pd
from sklearn.tree import DecisionTreeClassifier # Import Decision Tree Classifier
from sklearn.model_selection import train_test_split # Import train_test_split function
from sklearn.metrics import accuracy_score, precision_recall_fscore_support, confusion_matrix
#Libraries related to Graph
from six import StringIO
from sklearn.tree import export_graphviz
import pydotplus

#Step 2: Load the Dataset (file name : diabetes.csv)
col_names = ['nopreg', 'glucose', 'bp', 'skin', 'insulin', 'bmi', 'pedigree', 'age', 'label']
df = pd.read_csv("diabetes.csv", header=None, names = col_names)

#Step 3: Split Dataset in features and target variable
feature_cols = ['nopreg', 'glucose', 'bp', 'skin', 'insulin', 'bmi', 'pedigree', 'age']
X = df[feature_cols].values
y = df['label']
#print(X[0:5])

#Step 4 : Split in to Training and Test Data (20% of the data for testing)
X_train, X_test,y_train,y_test = train_test_split(X,y,test_size=0.2,shuffle=False, random_state=0)

#Step 5 : Build a model (Use DecisionTreeClassifier (criterion : Entropy, max_depth:3))
#3 Substeps - Create the model, Train the model, Predict response for test dataset
clf = DecisionTreeClassifier(criterion='entropy', max_depth=3)
```

```
clf.fit(X_train,y_train)
y_pred = clf.predict(X_test)

#Step 6 : Evaluate the model
#Compute Accuracy, Precision, Recall, F1-Score, Support and Confusion Matrix
#Round of Accuracy, Precision, Recall and F1-Score to 2 decimal precision using round function
accuracy = accuracy_score(y_test,y_pred)
print("Accuracy : ", round(accuracy,2))
prfs = precision_recall_fscore_support(y_test,y_pred,average='macro')
print("Precision : ", round(prfs[0],2))
print("Recall : ", round(prfs[1],2))
print("F1-Score : ", round(prfs[2],2))
print("Support : ", prfs[3])
print("Confusion Matrix")
print(confusion_matrix(y_test,y_pred))

#Step 7 : Test the Model with new data from the user
t1 = input('Enter input data\n')
t1_list = t1.split(",")
t1_list = list(map(float,t1_list))
yt1 = clf.predict([t1_list])
print(yt1)

#Step 8 : Visualizing Decision Trees

dot_data = StringIO()
export_graphviz(clf, out_file=dot_data,
                filled=True, rounded=True,
                special_characters=True, feature_names = feature_cols,class_names=['0','1'])
```

```
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
graph.write_png('diabetes.png')
```

4. Machine Learning Regression

Question 4: ML / Regression

```
#import required libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import math
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error,mean_squared_error

#Step 1 : Load the dataset
df = pd.read_csv("TSLA.csv")

#Step 2 : Shape or Dimensions of the DataFrame
print("DF Shape")
print(df.shape)
print()

#Step 3 : Explore Data Types of the various columns
print("DF Column Types")
print(df.dtypes)
print()

#Step 4 : Reindex data using a Datetime Index
df.set_index(pd.DatetimeIndex(df['Date']), inplace=True)

#Step 5 : Explore data by displaying 5 rows
```

```
print("Top 5 rows")
print(df.head())
print()

#Step 6 : Printing Statistical Summary of the dataset
print("Statistical Summary")
print(df.describe())
print()

#Step 7 :  Checking for empty data fields in the dataset
print("Columns and null values")
print(df.isnull().sum())
print()

#Step 8: Split dataset in features and target variable
X = df[["Open","High","Low","Volume"]]
y = df["Close"]

#Step 9 : Split dataset into training set and test set (30% test data size)
X_train, X_test, Y_train, Y_test = train_test_split(X,y,test_size=0.3, shuffle=False, random_state=0)

#Step 10 : Create the Linear Regression model
lr = LinearRegression()

#Step 11 : Train the Linear Regression Model
lr.fit(X_train, Y_train)

#Step 12 : Predict the response for test dataset
Y_test_pred = lr.predict(X_test)
```

```
#Step 13 : Create a new DataFrame with the Actual Price and Predicted Price for the test dataset
#Print the first 20 values
dfr = pd.DataFrame({"Actual_Price" : Y_test, "Predicted_Price" : Y_test_pred})
print("First 20 values")
print(dfr.head(20))
print()

#Step 14 : Generate line plots for Actual Price and Predicted Price
#Plot Title : Regression Analysis, xlabel : Date, ylabel : Actual/Predicted Price, Title : Regression Analysis
sns.lineplot(x = "Date", y = "Actual_Price", data = dfr)
sns.lineplot(x = "Date", y = "Predicted_Price", data = dfr)
plt.suptitle("Regression Analysis")
plt.xlabel("Date")
plt.ylabel("Actual/Predicted Price")
plt.savefig("Regression.png")


#Step 15 : Check the accuracy of the Linear Regression Model
print("Linear Regression Score")
print(round(lr.score(X_test, Y_test),2))
print()

#Step 16 : Print intercepts and coefficients
print("Intercept")
print(round(lr.intercept_,2))
print("Coefficients")
for coeff in lr.coef_:
   print(round(coeff,2))
```

```
print()

#Step 17 : Print Regression Evaluation Metrics
mae = mean_absolute_error(Y_test,Y_test_pred)
#squared True returns MSE value, False returns RMSE value.
mse = mean_squared_error(Y_test,Y_test_pred) #default=True
rmse = math.sqrt(mse)

print("MAE")
print(round(mae,2))
print("MSE")
print(round(mse,2))
print("RMSE:")
print(round(rmse,2))
```

5. **Natural Language Processing**

.

Question 5: Natural language processing

```
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
documents=input("Enter a sentence: ")
words=word_tokenize(documents)
words=[word for word in words if word  not in stopwords.words("english")]
documents=" ".join(words)
print(documents)
```

Key Takeaways:

- Better business decisions.
- Easier trend analysis and exploration.
- Acquire the dataset, Import all the crucial libraries & Import the dataset.
- Identifying and handling the missing values.
- Encoding the categorical data.
- Understanding Machine Learning. Machine learning is a subset of artificial intelligence that involves training machines to learn from data, without being explicitly programmed.
- Natural Language Processing

Conclusion:

- Data preprocessing It is the conversion of the data to useful information. Broadly, data processing involves six basic steps: data collection, data storage, data sorting, data processing, data analysis, and data presentation, leading to conclusions.

- Machine learning (ML) refers to a system's ability to acquire, and integrate knowledge through large-scale observations, and to improve, and extend itself by learning new knowledge rather than by being programmed with that knowledge.

SUDHARSAN ENGINEERING COLLEGE

Sathiyamangalam, Pudukkottai.

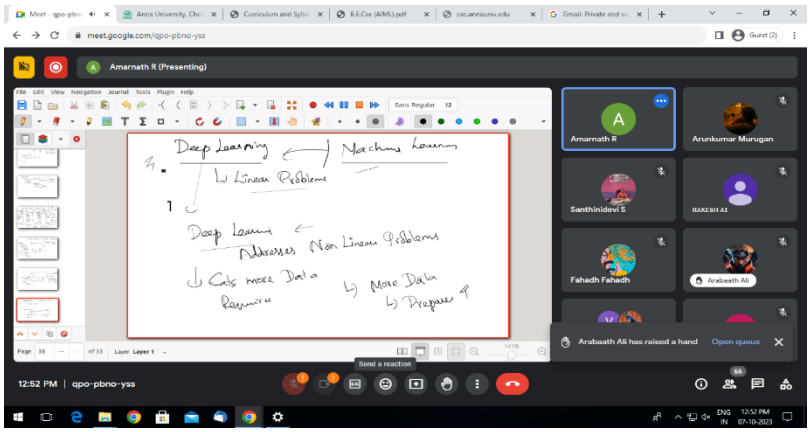
One Day National Level Workshop
on
"Tech Skills in AI Era and Job Opportunities"

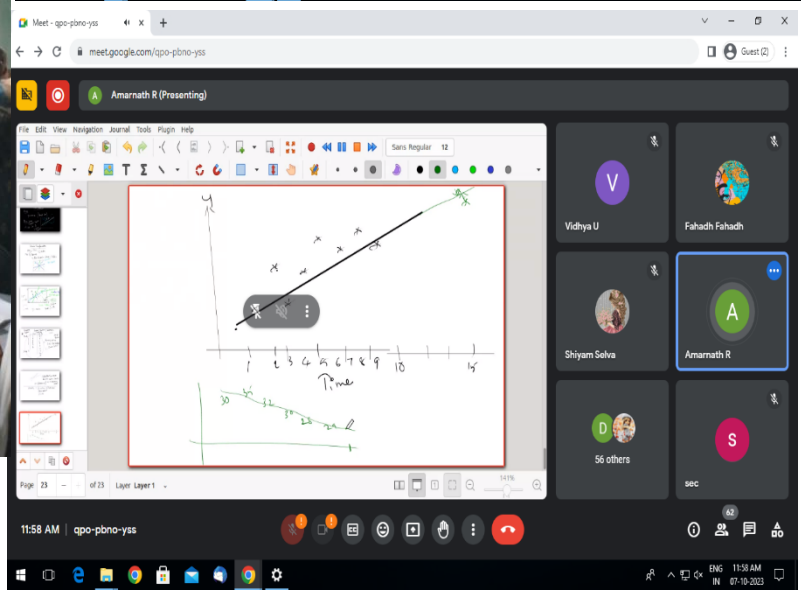
Report

PARTICIPATION CERTIFICATE

One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

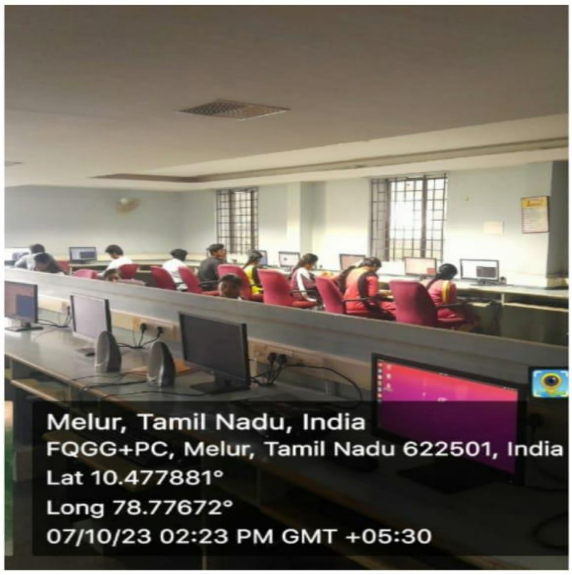
Session 1 - Photos



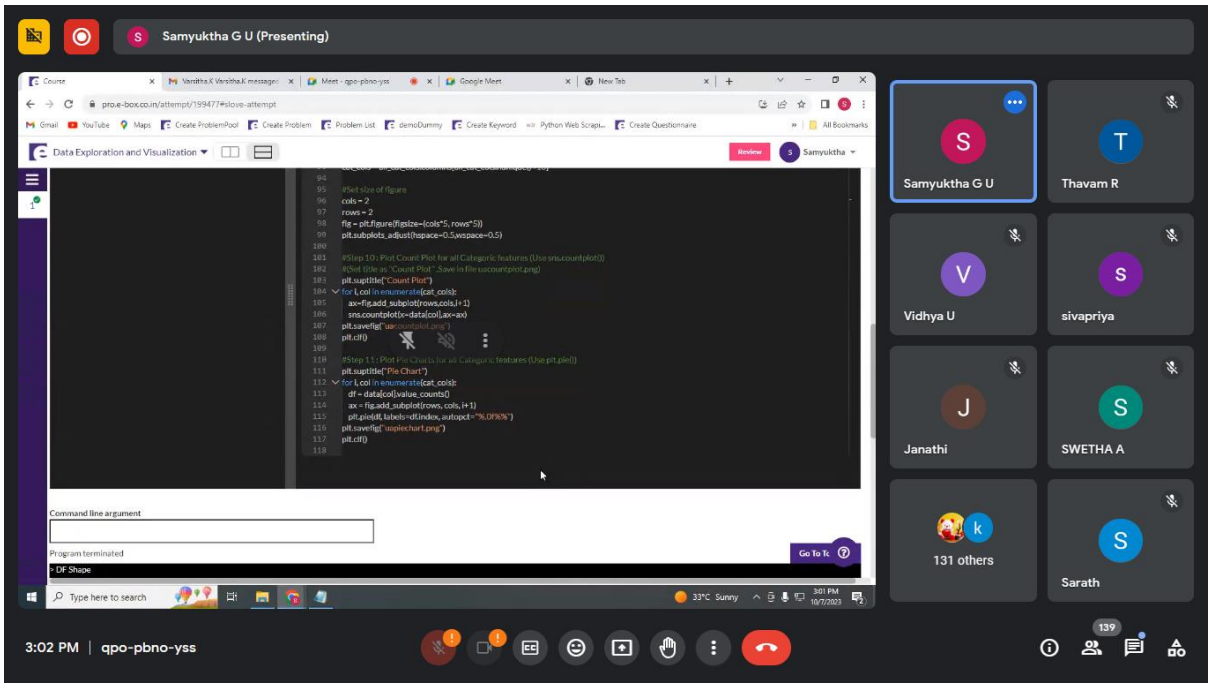


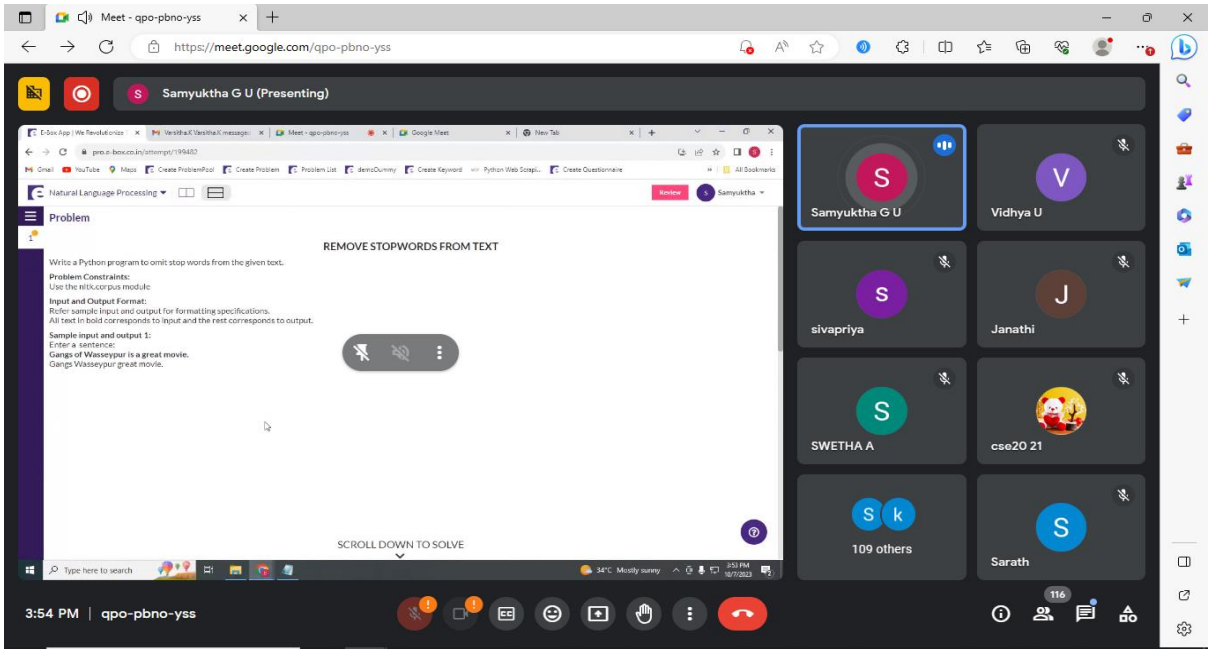
One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

Session 2 - Photos



Session 2 - Photos





One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

Feed Back Form

Name: *

A SARATH SARATHY

E-mail: *

sarathsarathy79@gmail.com

1. The workshop's content was well-organized and easy to follow *

- ○ Strongly Agree
- ◉ Agree
- ○ Neutral
- ○ Strongly Disagree
- ○ Disagree

2.The workshop inspired me to pursue further learning in AI and tech skills *

- ○ Strongly Agree
- ◉ Agree
- ○ Neutral
- ○ Strongly Disagree
- ○ Disagree

9.Hands-on training is more effective for learning tech skills than theoretical instruction *

- (●) Strongly agree
- ( ) Agree
- ( ) Neutral
- ( ) Strongly Disagree
- ( ) Disagree

10.How satisfied are you with the interactive elements and engagement during the virtual workshop? *

- ( ) 5. Excellent
- (●) 4.Very Good
- ( ) 3. Good
- ( ) 2. Satisfactory
- ( ) 1.Very Poor

Submit Clear form

Never submit passwords through Google Forms.

This form was created inside of Sudharsan Engineering College. Report Abuse

Google Forms

One Day National Level Workshop on "Tech Skills in AI Era and Job Opportunities"

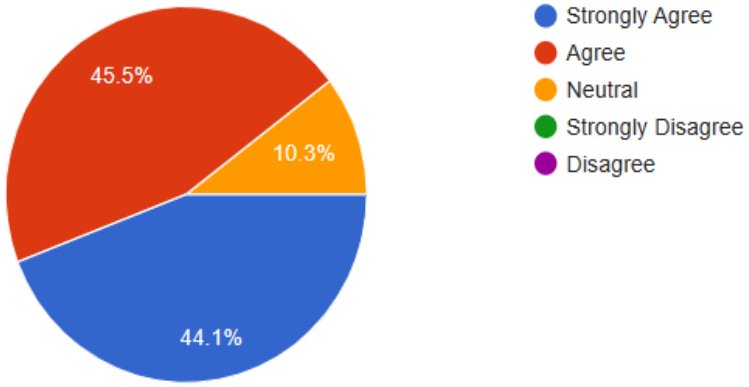
Feedback Analysis Report

Name:	1. The workshop's content was well-organized and easy to follow	2.The workshop inspired me to pursue further learning in AI and tech skills	3.The workshop effectively addressed the challenges and obstacles in the AI field	4.The presenter(s) demonstrated a strong knowledge of the subject matter	5.The workshop effectively showcased the importance of AI in various industries	6.The workshop included diverse perspectives and voices in the tech industry	7.Overall, I believe this workshop added value to my knowledge and skills	8.The workshop was well-timed and didn't run too long.	9.Hands-on training is more effective for learning tech skills than theoretical instruction	10.How satisfied are you with the interactive elements and engagement during the virtual workshop?
MARUDHAN T	Agree	Agree	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Neutral	Agree	4.Very Good
NAGARAJAN.N	Strongly Agree	Strongly Agree	Agree	Agree	Neutral	Agree	Agree	Agree	Agree	5. Excellent
V.RAMYA	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly agree	5. Excellent
DEIVANAI R	Strongly Agree	Agree	Agree	Agree	Neutral	Strongly Agree	Strongly Agree	Agree	Agree	3. Good
ARAVINTH N	Strongly Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	5. Excellent
VIDHYA U	Agree	Agree	Agree	Agree	Agree	Agree	Strongly Agree	Agree	Agree	4.Very Good
VISVESHVARAN M	Agree	Strongly Agree	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Strongly agree	5. Excellent
MOHAMED HANIFA B	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Agree	Strongly agree	5. Excellent
P.SRILEGA	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Disagree	Strongly agree	5. Excellent
RITHIK ROSHAN B	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly agree	4.Very Good

1. The workshop's content was well-organized and easy to follow

Copy

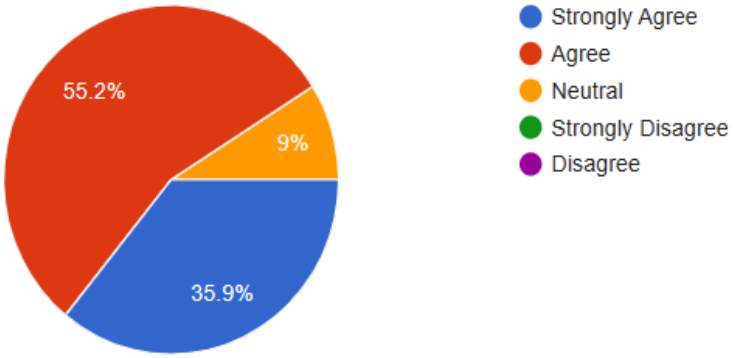
145 responses



2.The workshop inspired me to pursue further learning in AI and tech skills

Copy

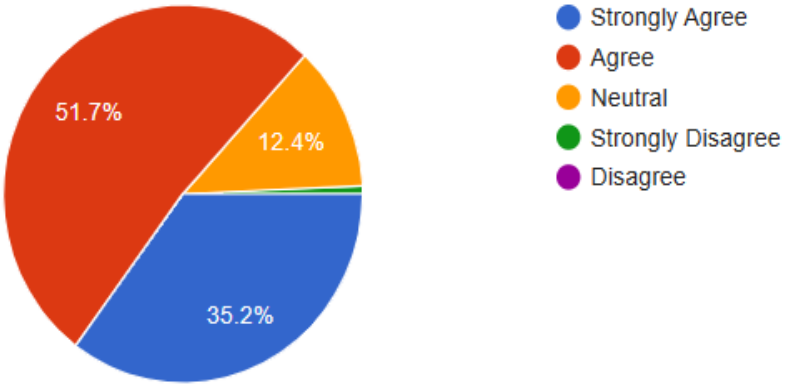
145 responses



3.The workshop effectively addressed the challenges and obstacles in the AI field

Copy

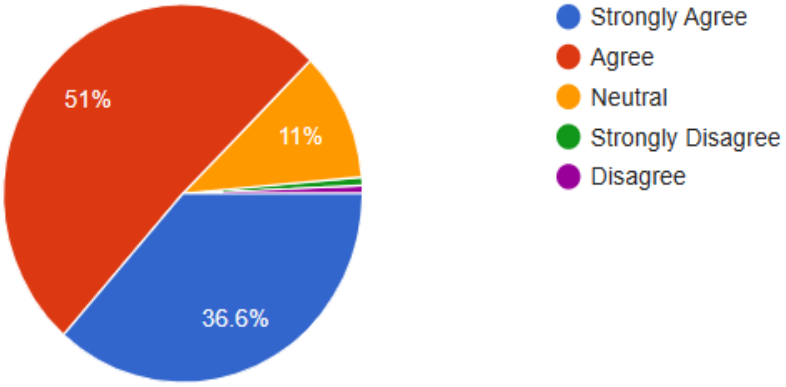
145 responses



4.The presenter(s) demonstrated a strong knowledge of the subject matter

Copy

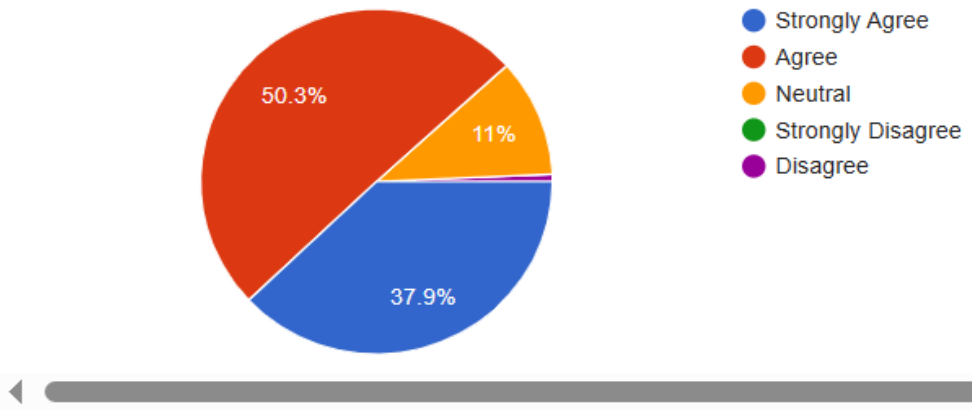
145 responses



5.The workshop effectively showcased the importance of AI in various industries

Copy

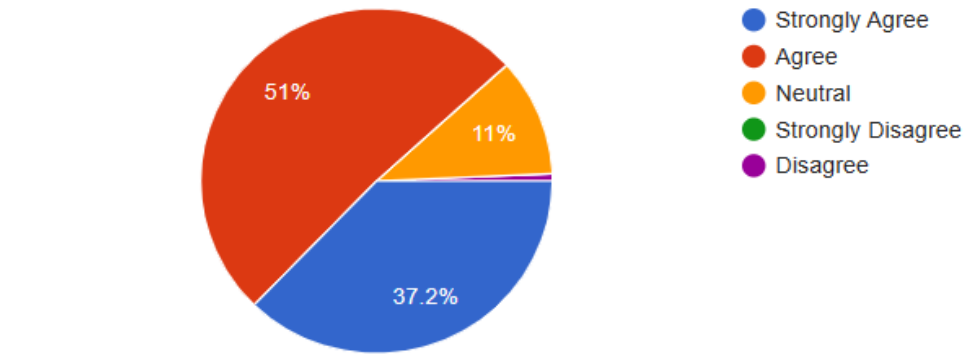
145 responses



6.The workshop included diverse perspectives and voices in the tech industry

Copy

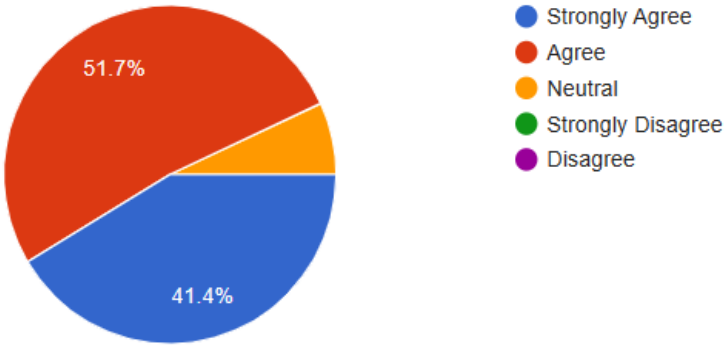
145 responses



7.Overall, I believe this workshop added value to my knowledge and skills

Copy

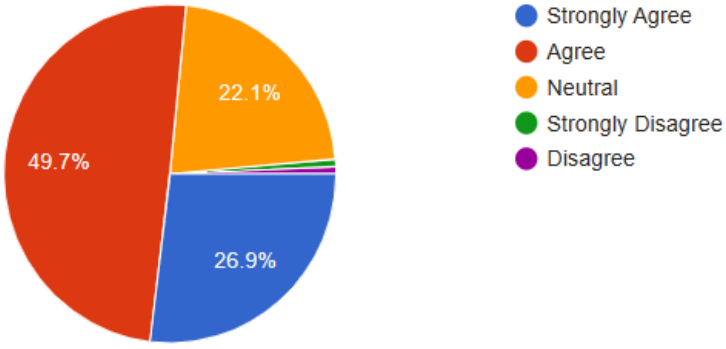
145 responses



8.The workshop was well-timed and didn't run too long.

Copy

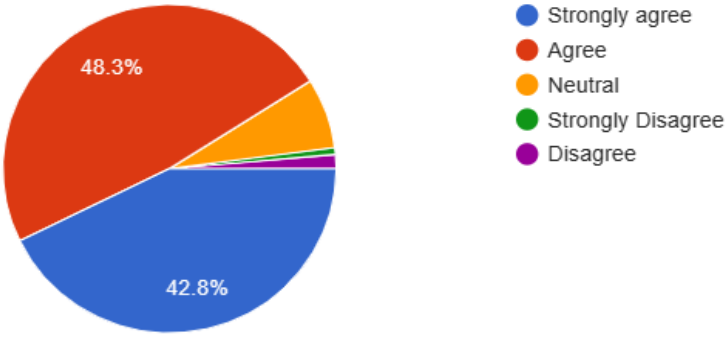
145 responses



9.Hands-on training is more effective for learning tech skills than theoretical instruction

Copy

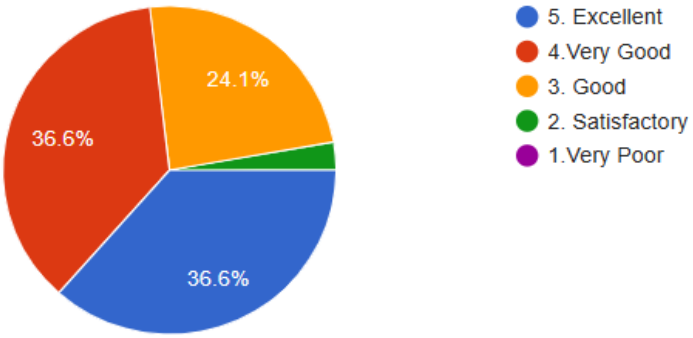
145 responses



10.How satisfied are you with the interactive elements and engagement during the virtual workshop?

Copy

145 responses



Introduction of E-box Platform usage to the Freshers - 07.08.2023