

STUDENT INTERNSHIP PROGRAM (SIP) REPORT

CELEBAL TECHNOLOGIES

Submitted By

Harshal Santosh Tawade

Roll No – 04

PRN No. - 0120200027

SCHOOL OF COMPUTER ENGINEERING

MIT ACADEMY OF ENGINEERING

ALANDI (D), PUNE

CERTIFICATE

This is to certify that the “**Student Internship Program (SIP)**” Report submitted by **Harshal Santosh Tawade**, PRN No.- **0120200027** is work done by him and is submitted during the **2023-24** academic year.

Faculty Mentor

Mrs. Vinodini Gupta

School - Internship Coordinator

Mr. Nilesh Navghare

School Dean

Dr. Rajeshwari Goudar

Internship Certificate



CERTIFICATE OF INTERNSHIP

This Certificate is Proudly awarded to :

Harshal Santosh Tawade

in recognition of your outstanding completion of **Celebal Summer Internship Program** at Celebal Technologies under **Data Science** domain for the duration of **10th June 2023 to 8th August 2023**.

You are found to be hardworking, enthusiastic and diligent during your internship. We highly appreciate your accomplishments and wish you the best for the future.

A handwritten signature in black ink, appearing to read 'Sharthak'.

Sharthak Acharjee
Senior Manager - HR



ACKNOWLEDGEMENT

I am immensely grateful to the entire Celebal Technologies team for providing me with an exceptional opportunity to undertake my internship at their esteemed organization. Their guidance, support, and insights have played a pivotal role in shaping my understanding of the dynamic field of Data Science.

I extend my heartfelt gratitude to Mr. Rahul Singh Rathore, my mentor and team lead, for his invaluable guidance, continuous encouragement, and wealth of knowledge that he shared with me throughout the internship. His mentorship has been instrumental in enhancing my technical skills and professional growth. I would like to thank the entire Celebal team. Their eagerness to work together, passion, and willingness to impart their knowledge have produced a vibrant learning atmosphere that has really enhanced my experience.

On the academic front, I am indebted to Mrs. Vinodini Gupta , my guide and mentor from college. Her insightful feedback, thoughtful discussions, and encouragement were instrumental in shaping the direction of my internship project.

I extend my sincere appreciation to Dr. Rajeshwari Goudar, Dean of the School of Computer Engineering, Mr. Nilesh Navghare, SIP Coordinator at MIT Academy of Engineering, for granting me the permission to undertake this internship. Their continuous support, guidance, and encouragement have been a driving force behind my pursuit of this endeavor.

In conclusion, I am grateful for the collective support of the Celebal Technologies team, my academic mentors, and the college administration. This internship experience would not have been possible without their unwavering guidance and encouragement.

Harshal Santosh Tawade

Table of Contents

Sr. NO.	TOPIC	Page No.
1	Introduction	
	1.1 About the Organization	06
	1.2 Timeline	11
2.	Internship Discussion	
	2.1 Skilling Program	14
	2.2 Python Basics	14
	2.3 Power BI Tool	16
	2.4 Projects	17
	2.5 Skills	20
	2.6 Challenges	21
	2.7 Outcomes	23
3.	Conclusion	24
4.	References	25

1. Introduction

1.1 About the Organization –

Celebal Technologies is a premier software services company in the field of Data Science, Big Data, and Enterprise Cloud. They help organizations achieve a competitive advantage with intelligent data solutions, built using cutting-edge technology. Their core offerings are around Cloud Innovation, Supply Chain Analytics, Chat Bots, Power Platforms, and Data Analytics.

Celebal Technologies was founded in 2012 by a team of experienced professionals with a passion for data and technology. They are headquartered in Pune, India, and have offices in Bangalore, Hyderabad, and Jaipur. The company has a team of over 500 professionals who are experts in a variety of technologies, including Microsoft Azure, AWS, Google Cloud Platform, Hadoop, Spark, and Hive.

Celebal Technologies has a proven track record of success in delivering innovative data solutions to clients in a variety of industries, including manufacturing, healthcare, financial services, and retail. They have been recognized by Microsoft as a Gold Partner and a Cloud Solution Provider.

Vision and Goals –

The vision of Celebal Technologies is to be a leading provider of data solutions that empower organizations to make better decisions. Their goals are to:

- Make data simple and easy to understand for all organizations.
- Provide solutions powered by modern cloud and artificial intelligence that integrate with traditional enterprise software.
- Help enterprises maximize productivity, and improve speed and accuracy.
- Be a trusted partner for organizations on their digital transformation journey.

Celebal Technologies is committed to these goals and is constantly innovating to provide the best possible data solutions to its clients. They believe that data is the key to unlocking new opportunities and improving the way organizations operate.

Initiatives

Here are some of the specific initiatives that Celebal Technologies is taking to achieve its vision and goals:

- Investing in research and development to stay ahead of the curve in data science and artificial intelligence.
- Building partnerships with leading technology companies to offer a wider range of solutions.
- Expanding its team of data experts to meet the growing demand for its services.
- Developing a strong focus on customer satisfaction to ensure that its clients are successful.

Celebal Technologies is well-positioned to achieve its vision and goals. With its strong track record, commitment to innovation, and focus on customer satisfaction, the company is poised to be a leading provider of data solutions in the years to come. Here are some of the key features of Celebal Technologies:

- Expertise in a wide range of data technologies
- Proven track record of success in delivering data solutions
- Strong focus on innovation
- Commitment to customer satisfaction

Impact

Celebal Technologies has had a significant impact on the field of data science and analytics. They have helped to make data more accessible and understandable for organizations of all sizes. Their solutions have helped businesses to improve their decision-making, optimize their operations, and gain a competitive advantage.

Here are some statistical data about Celebal Technologies

- The company has over 500 employees.
- The company has served over 100 clients in a variety of industries.
- The company has helped its clients to save over \$10 million in costs.
- The company has helped its clients to improve their decision-making by 20%.
- The company has helped its clients to increase their revenue by 15%.

Here are some of the awards and recognitions that Celebal Technologies has received:

- Microsoft Gold Partner
- Microsoft Cloud Solution Provider
- NASSCOM Top 50 Global Innovators
- Dataquest Top 25 Data Science Companies in India
- Analytics India Magazine Top 10 Data Science Companies in India

Here are some of the key areas where Celebal Technologies has had an impact:

- **Manufacturing:** Celebal Technologies has helped manufacturers to improve their production processes, reduce costs, and increase efficiency. They have done this by using data to identify areas where improvements can be made, such as by optimizing production lines or improving supply chain management.
- **Healthcare:** Celebal Technologies has helped healthcare organizations improve patient care, reduce costs, and improve operational efficiency. They have done this by using data to identify patients who are at risk for certain diseases, develop personalized treatment plans, and improve the efficiency of clinical trials.
- **Financial services:** Celebal Technologies has helped financial services organizations to detect fraud, improve risk management, and make better investment decisions. They have done this by using data to analyze customer transactions, identify patterns of suspicious activity, and develop models to predict future outcomes.
- **Retail:** Celebal Technologies has helped retailers to improve their customer experience, personalize marketing campaigns, and optimize inventory management. They have done this by using data to understand customer behavior, identify trends, and make better decisions about product assortment and pricing.

Commitment

Celebal Technologies is committed to making data simple and easy to understand for all organizations. They provide solutions powered by modern cloud and artificial intelligence that integrate with traditional enterprise software. Their tailor-made solutions help enterprises maximize productivity, improve speed and accuracy

1.2 Timeline -

1. Application Phase

The journey of my internship with Celebal Technologies began with the application phase. I submitted my application, expressing my interest in joining their dynamic team and showcasing my enthusiasm for Data Science and AI technologies.

2. Aptitude and Coding Test

Following the application phase, I was invited to participate in an aptitude and coding test. This phase evaluated my problem-solving skills, logical thinking, and coding proficiency. The test provided insights into my technical abilities and served as an initial screening process.

3. Offer Letter and HR interaction

Upon successfully clearing the aptitude and coding test, I was offered the internship. This intensive learning period was designed to enhance my skills in the Data Science and Machine Learning Field. Through HR interaction we were made familiar with the company culture and further process of the internship. Here we were assigned with our mentors.

4. Training Period – Week 1 to Week 4

In this training period, we were made familiar with the basics of concepts of data science and machine learning. Tasks were allocated weekly and feedback sessions were organized every Saturday in order to clear the doubts and queries of the mentor.

5. Project Allocation and Working

Upon successfully completing the training phase, I transitioned to the project allocation phase. During this stage, I was assigned a specific project that aligned with my skills, interests, and the organization's requirements. With the guidance of mentors and team members, I delved into the intricacies of the project, applying the knowledge and skills I had gained during the skilling program.

The subsequent weeks were characterized by hands-on work, problem-solving, collaboration, and continuous learning. Regular interactions with mentors and team members facilitated a seamless integration into the team's projects and tasks. This phase provided valuable practical experience and insight into the real-world applications of data science and machine learning technologies.

Throughout the internship timeline, I had the opportunity to engage with industry experts, work on cutting-edge projects, and contribute to the organization's goals. The structured progression from application to project

allocation ensured a comprehensive and enriching internship experience, allowing me to develop both technical and professional skills.

2. Internship Discussion

2.1 Skilling Program:

In **week 1**, I learned about the different data mining and data manipulation techniques. I applied these techniques to extract meaningful insights from data and clean and prepare data for analysis.

In **week 2**, I learned the basics of Python, including its syntax and data types. I also learned how to use Python libraries for data science, such as NumPy, Pandas, Matplotlib, and Scikit-Learn.

In **week 3**, I practiced with different machine learning algorithms. I learned how to choose the right machine-learning algorithm for a given problem, train and evaluate machine-learning models, and deploy machine-learning models in production.

In **week 4**, I learned about deep learning and its different techniques. I also learned how to apply deep learning techniques to solve real-world problems.

2.2 Python Basics

- **Data types**

- **Integers:** Integers are whole numbers, such as 1, 2, 3, and 4.
- **Floats:** Floats are decimal numbers, such as 1.0, 2.5, and 3.14.
- **Strings:** Strings are sequences of characters, such as "hello world".
- **Lists:** Lists are ordered collections of objects, such as [1, 2, 3, 4].
- **Tuples:** Tuples are immutable lists, such as (1, 2, 3, 4).
- **Dictionaries:** Dictionaries are mappings from keys to values, such as {"name": "John Doe", "age": 30}.

- **Operators**

- **Arithmetic operators:** Arithmetic operators are used to perform mathematical operations on numbers, such as +, -, *, /, and **.
- **Comparison operators:** Comparison operators are used to compare two values, such as ==, !=, <, >, <=, and >=.
- **Logical operators:** Logical operators are used to combine logical expressions, such as and, or, and not.
- **Membership operators:** Membership operators are used to test if a value is a member of a sequence, such as in and not in.

- **Control flow statements**

- If statements: If statements are used to execute code only if a certain condition is met.
- For loops: For loops are used to execute code repeatedly for a certain number of times.
- While loops: While loops are used to execute code repeatedly until a certain condition is met.

- **Functions**

- Functions are reusable blocks of code that can be called to perform a specific task.
- Modules: Modules are libraries of functions and classes that can be imported into Python programs.

2.3 Project Allocated

I was allocated a project that was Human Action Recognition based project in which the task was to produce the action of the given image using Deep Learning such as tensor flow, keras, preprocessing techniques and also deploy the model using Streamlit.

1) Image Classification using ML and DL model :

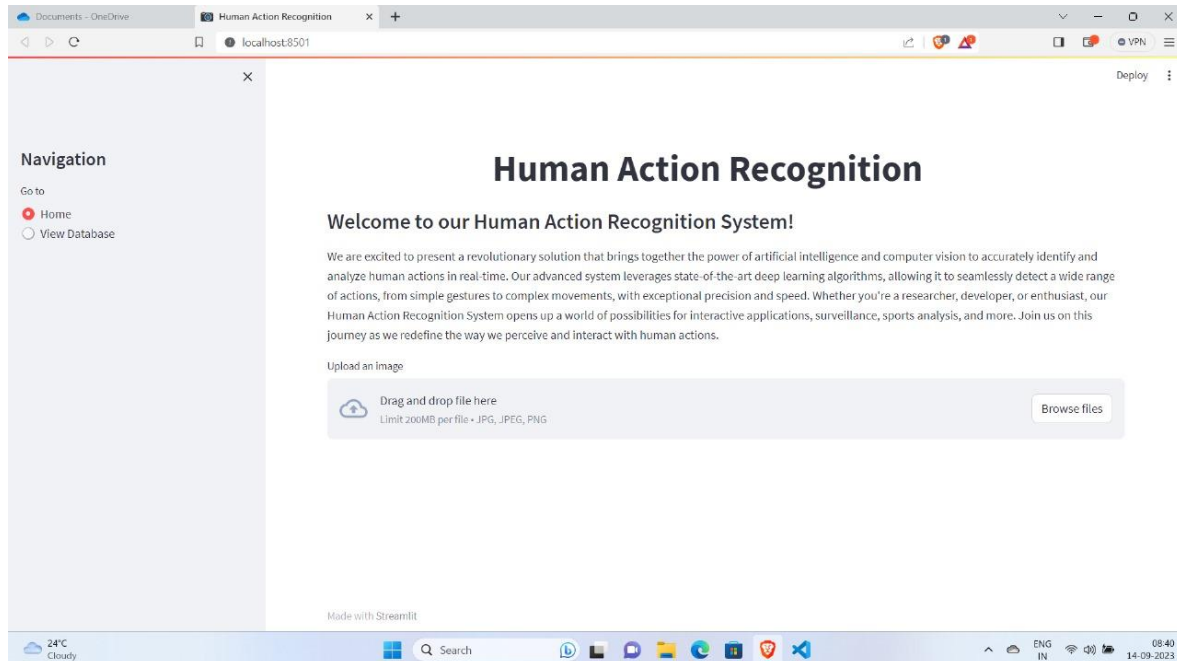
Tasks allocated:

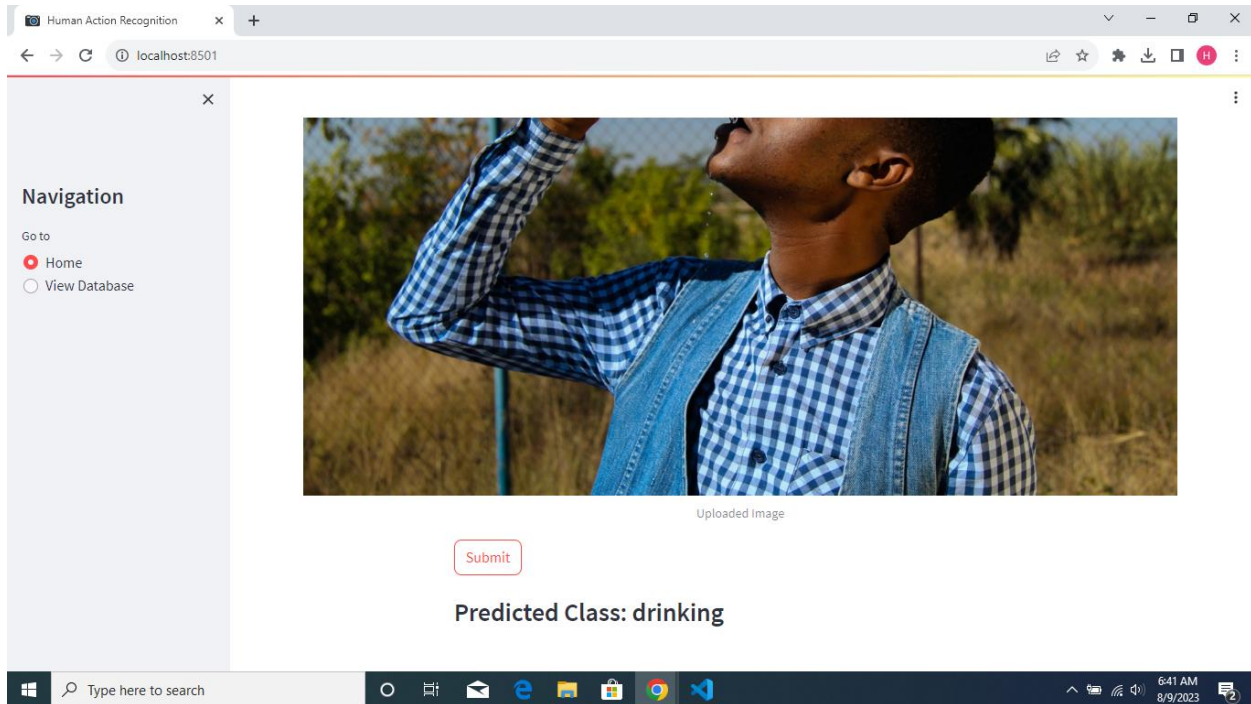
- Study about the ML and Deep Learning and understand the various libraries that are used in industry.
- Create a new virtual environment for the project so that the dependencies can be organized well.
- Study the in-built methods that are available in TensorFlow library, keras library, Gradient Class Activation Map for Heatmap.
- Work on Inception-ResNet-v2 which is trained on more than a million images.
- Implement the basic model help to classify the image in various class and utilize the different method for the web application.

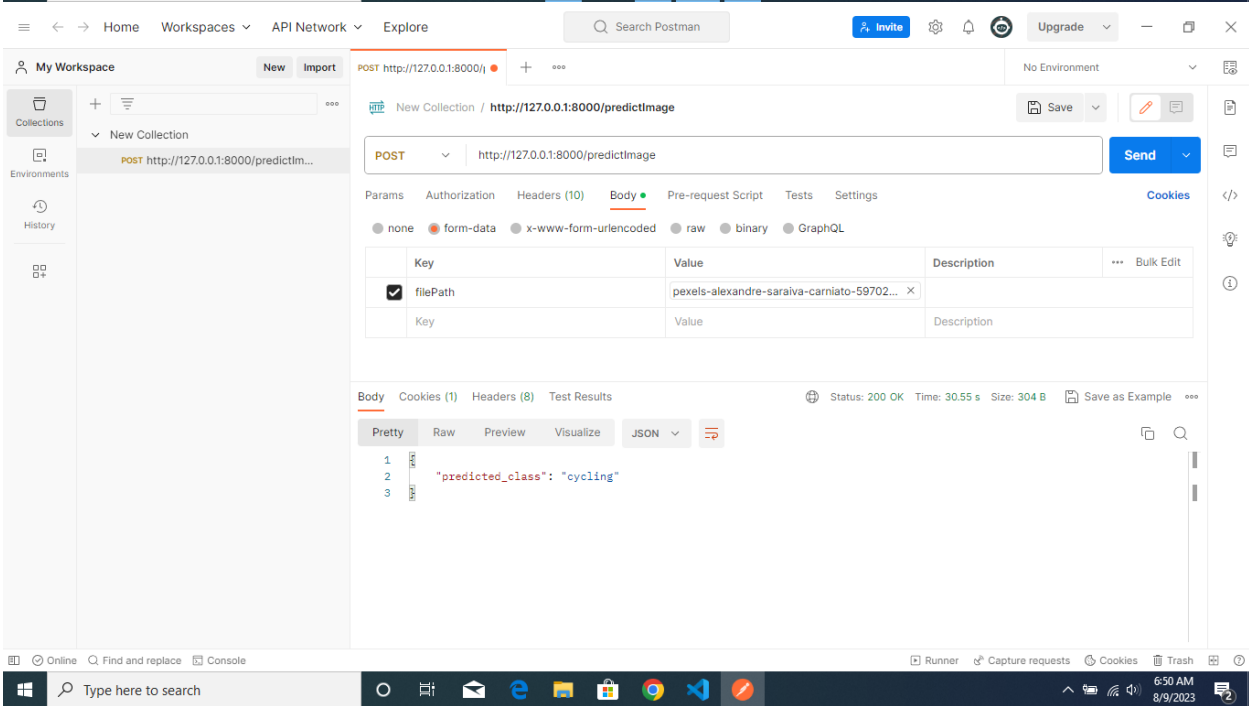
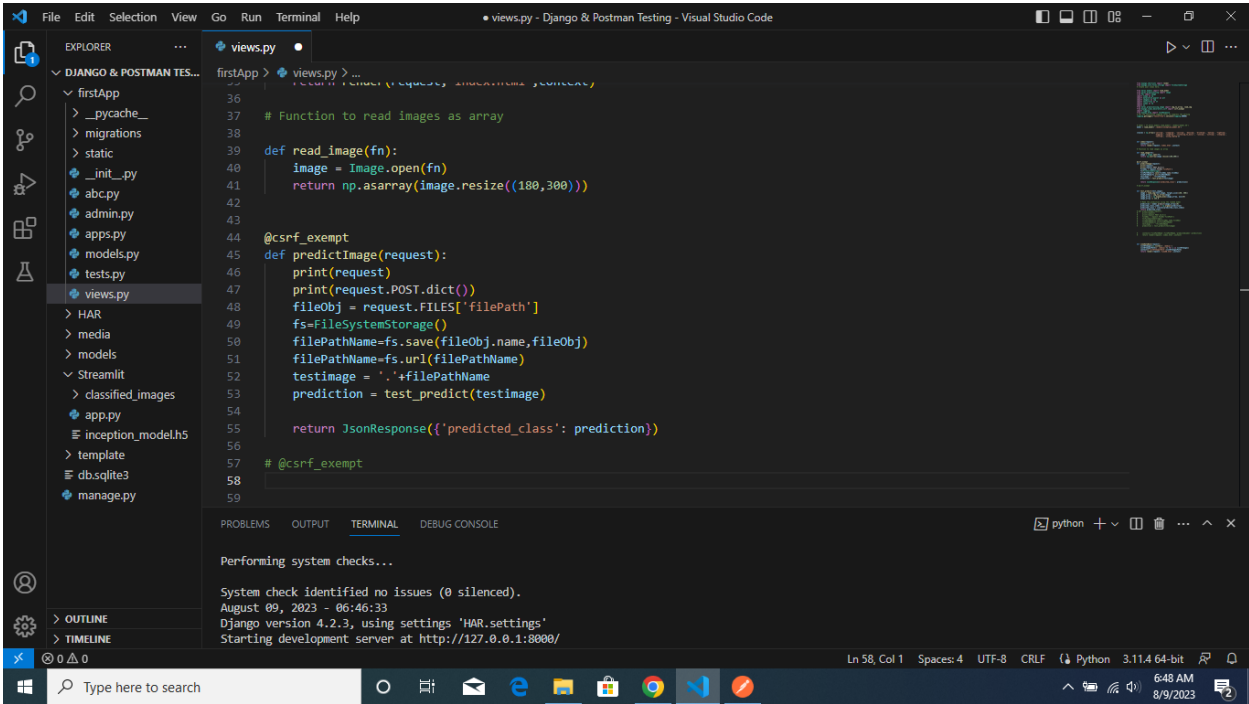
Task Performed:

- Designed basic function that train and test the image and outputs the class of image {i.e. calling', 'clapping', 'cycling', 'dancing', 'drinking', 'eating', 'fighting', 'hugging', 'laughing', 'listening to music', 'running' 'sitting', 'sleeping' 'texting', 'using laptop' }.
- Integrated the Human Action Recognition with the help of Streamlit created a user friendly web application so users can access the model.

User Interface of Web Application:







2.3 Skills Gained

- Technical skills:
 - Programming languages: Python ◦ Deep Learning model ,Preprocessing and TensorFlow ,keras Library ◦ Gradient Class Activation Map for Heatmap
 - DJANGO and Postman Testing
- Problem-solving skills:
 - Identifying and solving data science problems ◦ Asking questions, challenging assumptions, and evaluating evidence
- Communication skills:
 - Communicating your findings to stakeholders in a clear and concise way
 - Writing reports and explaining your findings to stakeholders
- Teamwork skills:
 - Working on projects with other people and collaborating with your colleagues
 - Asking for help when you need it
- Critical thinking skills:
 - Thinking critically about data and drawing conclusions ◦ Being open to new experiences and being willing to learn new things
- Adaptability skills:

- Being able to work under pressure and meet deadlines ○ Being flexible and willing to change your plans as needed

2.6 Challenges

- **Time management:** Data science projects can be time-consuming, so it was important to be able to manage my time effectively.
- **Troubleshooting:** Debugging and troubleshooting is the process in which I learned more about tools that are used in the project.

- Meetings :

Top Screenshot: Random Forest

Celebal Summer Internship | Data Science | Celebal Technologies | Rahul Singh Rathore | Batch -1

13:14

workshop_3.pdf | Supervised_Regre... | Classification Algo... | Cluster_kmean... | Artificial Neural N... | Convolutional Ne...

File | C:/Users/AashishJangid/OneDrive%20-%20Celebal%20Technologies%20Private%20Limited/Desktop/s.in/workshop_... | 15 of 35

numerical values based on splitting rules at each node of the tree.

Random Forest

Random Forest is an ensemble learning method that combines multiple individual decision trees to create a robust and accurate regression model

Basically, this create multiple decision tree and take average of all in last.

It use ensemble learning like - Bootstrap aggregation (bagging)
 A training dataset convert in many subset by including random rows (include repetition) and subset features that

X dataset

N_1 feature N_2 feature N_3 feature N_4 feature

Class 0 Class 0 Class 0 Class 0

MAJORITY VOTING

FINAL CLASS

Aashish Jangid (External) search 33°C 14:23 01-07-2023

Type here to search

Take control Pop out Chat People Raise React View More Camera Mic Share Leave

1/3

AJ 47

Aashish Jangid ... Participants

Bottom Screenshot: Unsupervised Learning and Clustering

Celebal Summer Internship | Data Science | Celebal Technologies | Rahul Singh Rathore | Batch -1

24:23

workshop_3.pdf | Supervised_Regre... | Classification Algo... | Cluster_kmean... | Artificial Neural N... | Convolutional Ne...

File | C:/Users/AashishJangid/OneDrive%20-%20Celebal%20Technologies%20Private%20Limited/Desktop/s.in/workshop_... | 21 of 35

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.

Unsupervised learning algorithms explore the data and identify inherent patterns or groupings without prior knowledge or guidance. They can be used for various purposes such as data exploration, data preprocessing, feature extraction, anomaly detection, and clustering.

Clustering: Clustering algorithms group similar data points together based on their characteristics or proximity in the feature space. The goal is to identify natural clusters or subgroups within the data. Examples of clustering algorithms include k-means clustering, hierarchical clustering, and DBSCAN.

STRUCTURED UNSTRUCTURED ALL DATA MODELS BUSINESS SYSTEMS

INPUT RAW DATA

Unlabeled data

Algorithm

Interpretation

Processing

OUTPUT

Dogs

Cats

Aashish Jangid (External) search 33°C 14:34 01-07-2023

Type here to search

Take control Pop out Chat People Raise React View More Camera Mic Share Leave

1/2

AJ 49

Aashish Jangid ... Participants

Conclusion

It was a wonderful experience working as a Data Science Intern at Celebal Technologies. For my continuous efforts and timely submission of project and overall performance I was awarded with the Silver Badge by the company

References

- <https://www.sciencedirect.com/science/article/pii/S2667096821000392>
- <https://www.worldscientific.com/doi/10.1142/S1469026821500267>
- <https://www.ijrar.org/papers/IJRAR2001915.pdf>
- https://link.springer.com/chapter/10.1007/978-3-642-25446-8_4
- <https://pyimagesearch.com/2019/11/25/human-activity-recognition-with-opencv-and-deep-learning/>