Industrial Training

on

Image Classification

SUBMITTED BY

Mayank Kumar CSE(AI&ML) -A

210962114 Mayankkumar02002@gmail.com

33 7851920834

Under the Guidance of:

Mr.Kunal Gautum

Director

Move37 Capital Limited

Diagram

Description automatically generated with medium confidence

Department of Computer Science and Engineering

July 2024

**I**

Move37 Capital Limited

Room 26 - Sbc House, Restmor Way, Wallington, Surrey, United Kingdom, SM6 7AH

24” July 2024

TO WHOMSOEVER IT MAY CONCERN

This is to certify that Mr. Mayank Kumar, student of Manipal Institute of Technology has successfully completed his projects titled Image Processing and Classification of Custom Dataset at Move37 Capital Limited, under the guidance of Mr. Kunal Gautam, Director, Move37 Capital Limited**.**

Mayank worked with technologies like Keras, Torchvision, Transfomers and Kmeans. He also acquired invaluable knowledge on data processing, neural network design and model evaluation.

Documentation of the application in the project report is not feasible due to its confidential nature. The interns are not allowed to share the application outside the workplace.

This project is in partial fulfillment of the requirements for the award of the degree BTech Computer Science for the academic year 2024-25. The duration of the project was from 20\* May 2024 to 19• July 2024. We wish him all the best in his life and career.

A close up of a signature

Description automatically generatedfor Move37 Capital Limited.

Sameer Desai Operations Manager

***II***

***ACKNOWLEDGEMENTS***

I am extremely grateful for the enriching experience I have had during my two-month summer internship at Move37 Capital Limited. This journey has been nothing short of transformative, and I owe my gratitude to a number of individuals and the supportive environment at Move37 Capital.

First and foremost, I would like to express my sincere appreciation to my mentor and supervisor, Kunal Sir. His guidance, patience, and willingness to share their expertise have been invaluable. His mentorship not only enhanced my technical skills but also helped me develop a deeper understanding of the industry. I am thankful his constructive feedback, which has undoubtedly contributed to my personal and professional growth.

I must also acknowledge the organizational support that made my internship seamless. The administrative staff, and everyone involved in the coordination of this internship deserve recognition for their efforts.

Lastly, I want to express my gratitude to my family and friends for their unwavering support throughout this journey. Their encouragement and belief in me motivated me to give my best.

As I conclude my internship and reflect upon this experience, I am leaving with a wealth of knowledge, memorable experiences, and a network of wonderful people. My time at Move37 Capital Limited has been instrumental in shaping my aspirations and has further fueled my passion for technology and innovation.

Thank you, Move37 Capital Limited, for this incredible opportunity. I look forward to the future with newfound confidence and enthusiasm.

**III**

***ABSTRACT***

In my internship, I worked mainly on two major projects of image processing i.e. Image clustering and classification based on deep-learning techniques. The first task was to cluster the MNIST dataset (images of handwritten digits) with KMeans, after reducing its dimensionality using Principal Component Analysis (PCA). This strategy facilitated the discovery of structures and clusters in data, judged using metrics like Accuracy.

The second project was the classification of a custom dataset that is stored in images. npy and labels. npy files. Here, the image had to be handled by implementing self-attention layers on top of a neural network model that was created using Keras. The feature extraction is improved which means better classification accuracy due to the self-attention mechanism. My experience was invaluable in providing ideas on data preprocessing, neural network design and model evaluation, showing that advanced machine learning approaches are suitable for solving more involved image analyses than the classic ILSVRC-style problems

**IV**

***TABLE OF CONTENTS***

|  |  |
| --- | --- |
| **TOPIC** | **PG. NO.** |
| Acknowledgement | 3 |
| Abstract | 4 |
| Details of the Organization | 6 |
| Introduction | 7 |
| Methodology | 8-12 |
| Results | 13-16 |
| Conclusion | 17 |
| References | 18 |
|  |  |

.

***DETAILS OF THE ORGANIZATION – MOVE37 CAPITAL LIMITED***

The concept of self-attention has revolutionized the field of machine learning, particularly in natural language processing and computer vision. Self-attention mechanisms enable models to weigh the importance of different input elements, allowing them to capture long-range dependencies and intricate patterns in the data. This approach contrasts with traditional methods that rely on local features and sequential processing, such as convolutional neural networks and recurrent neural networks. By focusing on the relationships between all elements of the input simultaneously, self-attention models, like Transformers, can achieve remarkable performance in tasks such as language translation, image classification, and more. The flexibility and scalability of self-attention make it a powerful tool for handling large datasets and complex tasks. As research continues to advance, the applications of self-attention are expected to expand, offering new possibilities for innovations in various domains, including healthcare, finance, and autonomous systems.

***6***

***INTRODUCTION***

Artificial intelligence and machine learning have led to revolutions in the way that we deal with tasks like image classification and pattern recognition. When I was interning, two different but related projects were investigated by me which were based on recent techniques in this area. These helped me grow my technical skills and understand better how neural networks can be utilized to handle challenges encountered by imaging science concepts

However, the first project was based on MNIST dataset that is a well-known reference point in computer vision as this dataset contains images of 70,000 hand-written digits (0-9) each of size 28x28 pixels. Traditional approaches emphasize convolutional neural networks (CNNs) in classifying these digits; however, our project went against the tide. We decided to do clustering instead. We wanted to find out whether we could get any underlying patterns or similarities among these numbers by applying KMeans clustering to preprocessed MNIST images. This approach not only provided a new dimension to digit recognition but also highlighted the possibilities of using unsupervised learning methods in image analysis.

Another project was to classify a custom dataset stored in images.npy and labels.npy files. This dataset had unique challenges such as non-uniform image patches that required specific feature extraction techniques. Let’s discuss the addition of self-attention layers into the neural network architecture using the Keras framework herein below. Long-range dependencies are captured by self-attention mechanisms and they have been applied in sequential data; however, this technique is not well known in image analysis. So, we adjusted these layers to handle non-uniform patches extracted from the dataset so that our model can better understand intricate details such that it can improve overall classification accuracy.

**7**

***METHODOLOGY***

***DATASET DESCRIPTION***

MNIST Dataset:

So, the MNIST dataset is a big name in machine learning & computer vision. It's got 70,000 grayscale pics of handwritten numbers. Each pic is 28x28 pixels. This big group is split into two parts: 60,000 pics for training and 10,000 pics for testing. Each picture shows a number from 0 to 9. So, it's like a with 10 possible answers. The pixel values go from 0 to 255. That shows how dark each pixel is. For our project, we changed the pixel values from 0 to 1 to help our neural network learn better.

Custom Dataset (images.npy and labels.npy):

For the second project, we used a different dataset stored in two files: images.npy and labels.npy. The images.npy file has grayscale pics. Each pic is shaped like this: (16, 8, 1). This means every pic is 16 pixels tall and 8 pixels wide with one color channel since it's grayscale. The labels.npy file holds class labels for these pics, showing one of three classes. It's another multi-class puzzle! We had to use patch extraction for processing features in these images, some parts were tougher than others. To tackle this, we used self-attention mechanisms in our neural network model. We also split the dataset into training and testing parts to see how well our model works

***PREPROCESSING***

MNIST Dataset:

First, we got the MNIST data using Keras's mnist.load\_data() function. This is u training and testing sets separately. To make things run faster during training, we changed the pixel values from 0 to 255 to 0 to 1 by dividing by 255.0. After that, each image was given a single channel, changing their shape from (28, 28) to (28, 28, 1). We then used Principal Component Analysis (PCA) to cut down the number of dimensions but keep important parts. This turned the images into 50 main parts. Also, we made the data more varied and better for training by resizing, random flipping, random rotation, & random zooming.

**8**

Custom Dataset (images.npy and labels.npy)

For our custom dataset, we first loaded images & labels from .npy files with numpy.load(). These images started at a shape of (16, 8) and were reshaped to include one channel, making them (16, 8, 1). The class labels were turned into one-hot encoded vectors to help train a neural network with categories. We divided the data into training and testing sets: the first 6000 images/labels for training and the next 2000 for testing. Because this dataset is unique, we cut out non-uniform patches from each image using specific positions & sizes that were hardcoded. These patches were reshaped to get them ready for the neural network model. This way, all data fit right for training and checking how well it worked.

***MODEL ARCHITECTURE***

The model for the MNIST dataset kicks off with a data augmentation layer. This part handles normalization, resizing, and then on to random flipping, rotation & zoom on the images. After that, it chops the images into non-overlapping 6x6 patches. These patches get encoded with positional embeddings.

At the heart of the model, there are 8 transformer blocks. Each block has layer normalization and multi-head attention—also, some multi-layer perceptrons (MLP) thrown in with skip connections. Finally, the output gets another round of normalization. Then, it's flattened, given some dropout, and sent through an MLP before hitting a dense layer. This last layer figures out class probabilities for the 10 MNIST classes.

For a custom dataset, things start a bit different. Non-uniform patches get extracted from the images. A self-attention layer deals with these patches. Dense layers create query, key, and value vectors. Attention scores are worked out and used for weighted sums.

The output from this is flattened next. Then it goes into a dense layer with ReLU activation. It ends in a final dense layer that uses softmax activation to classify 3 classes. Both models pack in advanced neural network stuff, all tuned to fit what the specific dataset needs to classify effectively.

**9**

A diagram of a process

Description automatically generated

Flowchart of self-attention layers

***MODEL TRAINING***

So, when training the MNIST dataset model, we got the Vision Transformer (ViT) ready. We used the AdamW optimizer with a learning rate of 0.001 and weight decay set at 0.0001. The loss function was Sparse Categorical Crossentropy. We checked metrics like accuracy & top-5 accuracy. This model went through 10 epochs with a batch size of 256. We watched performance by splitting validation to 10%. We also added a checkpoint callback to save the best weights based on validation accuracy.

After training, we loaded the best weights and tested the model on the test set. It got high accuracy and top-5 accuracy. To see how it did, we made plots showing loss, accuracy, and top-5 accuracy over those epochs.

For another try with a custom dataset model, images were prepped by getting non-uniform patches and reshaping them for the self-attention mechanism. Here, the model was compiled using the Adam optimizer; categorical cross-entropy loss, and performance was tracked with accuracy.

Then, it trained for 10 epochs, but with a smaller batch size of 32, keeping an eye out for overfitting by using a validation split. Afterward, performance was checked on the test set with metrics like loss & accuracy noted down. The self-attention layers played a big role in classifying images well. The model showed it could deal with tricky data patterns effectively.

***10***

***MODEL EVALUATION***

For the MNIST dataset, first, we loaded up the best weights saved during the training process. This choice was based on validation accuracy. Next, we tested the model on test set, which had images it hadn't seen before. We looked at test accuracy & top-5 accuracy to see how well the model could classify digits. The final results? Pretty impressive! The Vision Transformer (ViT) model made a lot of correct predictions, proving it worked well. Also, we took a look at the training history by plotting loss, accuracy, and top-5 accuracy over time. These plots showed how the model learned and got better, plus how data augmentation and regularization helped.

When it came to the custom dataset model, we did a similar evaluation after training but focused on a separate test set. We measured loss and accuracy to see how good the model was at classifying images into three classes. The self-attention mechanism did a great job in capturing intricate patterns in the data—this showed in our metrics. Comparing predicted labels with actual ones helped us check precision & reliability. In short, the model learned to tell apart different classes effectively thanks to those self-attention layers boosting its classification skills.

**11**

***VISUALIZATION***

Representation assumed a significant part in understanding the presentation and conduct of the models prepared on the MNIST and custom datasets. For the MNIST model, the preparation history was plotted to show the patterns in misfortune, exactness, and top-5 precision over the ages. These plots uncovered the model's expectation to learn and adapt, representing how the model improved with every age and how well it summed up to the approval set. The misfortune plot demonstrated how the model's forecast blunders diminished after some time, while the exactness plots showed the model's rising ability to accurately order the digit pictures. The main 5 precision plot gave extra knowledge into the model's presentation by showing the level of situations where the right name was among the main five anticipated marks.

For the custom dataset model, perception included plotting the misfortune and precision during preparing and approval stages. These plots were essential for diagnosing potential overfitting or underfitting issues. The predictable decline in misfortune and expansion in precision on both preparation and approval sets showed the model's successful learning and speculation.

Moreover, the last assessment measurements were imagined to give a reasonable outline of the model's exhibition. Besides, for the KMeans bunching task on the MNIST dataset, a 2D PCA dissipate plot of the grouped information was made. This representation showed the bunches recognized by the KMeans calculation, considering a visual evaluation of how well the grouping lined up with the genuine digit classes. Each bunch was addressed with an alternate tone, making it simple to notice the dispersion and detachability of groups. This visual portrayal helped in understanding the bunching execution and gave an unmistakable method for assessing the viability of the PCA-diminished highlight space for grouping err ands.

**12**

***RESULTS***

***MNIST DATASET***

***i)* ACCURACY :**

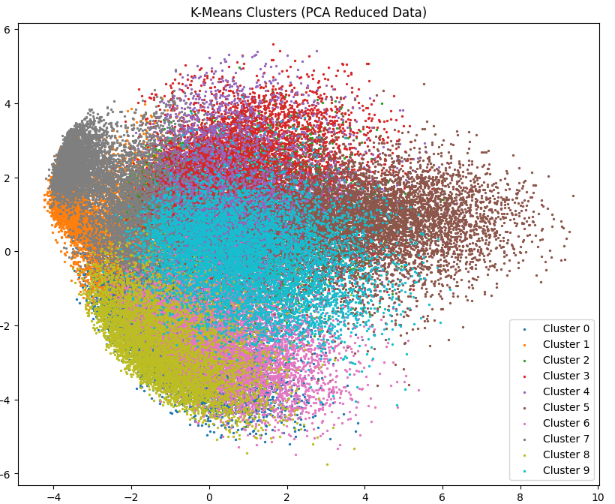
When we implemented K-means on MNIST Dataset using keras , we got an average accuracy of 58.46% as shown below

A screenshot of a computer

Description automatically generated

***ii)* CLUSTERS :**

There are 10 clusters which were presemt in MNIST Dataset, plotting these clusters we got the result as shown below:



**13**

***CUSTOM DATASET***

***i)* Epochs v/s Loss And Accuracy :**

Iterating over the dataset on a number of epochs , we get a improved accuracy and a decreased loss.

*A screenshot of a computer screen

Description automatically generated*

**ii) Plotting Model Loss :**

Plotting the model loss gives us a better result as shown below:

A graph showing the loss of a loss

Description automatically generated

**14**

**iii) Plotting Model Accuracy :**

Similar to loss , plotting the accuracy of model helps us to visualize the model better as shown :

A graph showing the growth of a model

Description automatically generated with medium confidence

**15**

**CONCLUSION**

In this task, we investigated the utilization of cutting edge AI methods to order and group digit pictures from the MNIST dataset and a custom dataset utilizing both traditional and best in class draws near. At first, we carried out a KMeans bunching calculation on the MNIST dataset to investigate solo learning methods. By diminishing the dimensionality of the information utilizing PCA, we accomplished critical bunching execution, with a Changed Rand Record (ARI) and Standardized Shared Data (NMI) scores that approved the viability of our methodology. Following the grouping task, we moved our concentration to managed picking up, utilizing the Vision Transformer (ViT) design for picture order. The ViT model, intended to deal with picture information by treating picture patches as arrangements, exhibited powerful execution on the MNIST dataset. The model accomplished high test precision and top-5 exactness, with perceptions of the preparation history highlighting its successful learning and speculation abilities. For the custom dataset gave as images.npy and labels.npy, we utilized a self-consideration system to order the pictures. The model's design was customized to catch the mind boggling conditions inside the information, bringing about noteworthy grouping precision on the test set. The preparation and approval plots affirmed the model's dependability and union, displaying the force of self-consideration layers in improving model execution. All through the undertaking, representation assumed a critical part in deciphering the outcomes and acquiring experiences into the model's way of behaving. The different plots, including misfortune and exactness patterns, gave an unmistakable story of the model's preparation process and its definitive viability. By and large, this venture exhibited the utility of both customary and current AI strategies in handling picture characterization and bunching errands. The outcomes highlight the significance of choosing proper models and methods in light of the particular idea of the dataset and the front and center concern

**16.**

**NBA/IET Mapping**

**NBA - PROGRAM OUTCOMES (PO) & PROGRAM SPECIFIC OUTCOMES (PSO)**

Engineering Graduates will be able to:

PO1:Engineering knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

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

PO3:Design/development of solutions: Design solutions for complex engineering 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 engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6:The engineer 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 engineering practice.

PO7:Environment and sustainability: Understand the impact of the professional engineering 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 engineering 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 engineering 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.

PO11: Project management and finance: Demonstrate knowledge and understanding of the engineering 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.

**17**

PSO1: Analyse and solve real world problems by applying a combination of hardware and software. PSO2: Formulate & build optimised solutions for systems level software & computationally intensive applications.

PSO3: Design & model applications for various domains using standard software engineering practices. PSO4: Design & develop solutions for distributed processing & communication.

**NBA CO PO Mapping**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CSE 4298-ITR** | **CO** | **PO1** | **PO2** | **PO3** | **PO4** | **PO5** | **PO6** | **PO7** | **PO8** | **PO9** | **PO10** | **PO11** | **PO12** | **PSO1** | **PSO2** | **PSO3** | **PSO4** |
| CSE  4298.1 | Understand the functioning of the industry | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| CSE  4298.2 | Understand the requirements of real world applications | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 0 | 3 | 2 | 0 | 1 | 1 | 1 | 2 | 0 |
| CSE  4298.3 | Demonstrate skills to use modern engineering tools, software and equipment to analyze problems | 0 | 0 | 2 | 2 | 3 | 1 | 1 | 0 | 3 | 2 | 0 | 1 | 1 | 1 | 2 | 0 |
| CSE  4298.4 | Demonstrate an ability to envisage and work on laboratory and multidisciplinary tasks | 0 | 0 | 2 | 1 | 3 | 1 | 1 | 0 | 3 | 2 | 0 | 1 | 1 | 1 | 2 | 0 |
| **CSE 4298 (Avg. correlation level)** | | **0.75** | **0.25** | **1.75** | **1** | **1.75** | **0.75** | **0.75** | **0.5** | **2.75** | **1.5** | **0** | **0.75** | **1.0** | **0.75** | **1.75** | **0** |

NBA Program articulation Matrix

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Course Code | Course Title | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 | PO10 | PO11 | PO12 | PSO1 | PSO2 | PSO3 | PSO4 |
| CSE 4298 | Industrial Training | 0.75 | 0.25 | 1.75 | 1 | 1.75 | 0.75 | 0.75 | 0.5 | 2.75 | 1.5 | 0 | 0.75 | 1.0 | 0.75 | 1.75 | 0 |

**18**

**IET-AHEP Learning outcome statements.**

|  |  |
| --- | --- |
| **IET-AHEP Learning outcome statements** | |
| ***Code*** | ***Learning outcome (LO)*** |
| C1. | Apply knowledge of mathematics, statistics, natural science and engineering principles to the solution of complex problems. Some of the knowledge will be at the forefront of the particular subject of study |
| C2. | Analyse complex problems to reach substantiated conclusions using first principles of mathematics, statistics, natural science and engineering principles |
| C3. | Select and apply appropriate computational and analytical techniques to model complex problems, recognising the limitations of the techniques employed |
| C4. | Select and evaluate technical literature and other sources of information to address complex problems |
| C5. | Design solutions for complex problems that meet a combination of societal, user, business and customer needs as appropriate. This will involve consideration of applicable health & safety, diversity, inclusion, cultural, societal, environmental and commercial matters, codes of practice and industry standards |
| C6. | Apply an integrated or systems approach to the solution of complex problems |
| C7. | Evaluate the environmental and societal impact of solutions to complex problems and minimise adverse impacts |
| C8. | Identify and analyse ethical concerns and make reasoned ethical choices informed by professional codes of conduct |
| C9. | Use a risk management process to identify, evaluate and mitigate risks (the effects of uncertainty) associated with a particular project or activity |
| C10. | Adopt a holistic and proportionate approach to the mitigation of security risks |
| C11. | Adopt an inclusive approach to engineering practice and recognise the responsibilities, benefits and importance of supporting equality, diversity and inclusion |
| C12. | Use practical laboratory and workshop skills to investigate complex problems |
| C13. | Select and apply appropriate materials, equipment, engineering technologies and processes, recognising their limitations |
| C14. | Discuss the role of quality management systems and continuous improvement in the context of complex problems |
| C15. | Apply knowledge of engineering management principles, commercial context, project and change management, and relevant legal matters including intellectual property rights |
| C16. | Function effectively as an individual, and as a member or leader of a team |
| C17. | Communicate effectively on complex engineering matters with technical and non-technical audiences |
| C18. | Plan and record self-learning and development as the foundation for lifelong learning/CPD |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **VII SEMESTER - CSE 4298 – Industrial Training - CLO - AHEP LO Mapping** | | | | | | | | | | | |
| **Course Learning Outcome** | **Statement** | **C1** | **C2** | **C3** | **C4** | **C6** | **C7** | **C8** | **C11** | **C15** | **C18** |
| **CLO 4298.1** | Apply mathematics, science and engineering skills to identify, formulate, synthesize and solve the problems from various areas of computer science engineering. | ü |  | ü |  |  |  |  |  |  |  |
| **CLO 4298.2** | Have knowledge of new trends in engineering/technology by developing programmable coding in various computer programming languages. |  |  |  | ü |  |  |  |  |  |  |
| **CLO 4298.3** | Use the industry standard tools to analyze, design, develop and test software engineering based applications. |  |  | ü |  |  |  |  |  |  |  |
| **CLO 4298.4** | Apply theoretical knowledge to real-world engineering problems and manage complex engineering projects. |  | ü |  |  | ü |  |  |  |  |  |
| **CLO 4298.5** | Understand the adverse societal impacts of the solutions to complex problems during project development. |  |  |  |  |  | ü |  |  |  |  |
| **CLO 4298.6** | Identify and analyze ethical concerns to be followed during the practice school/research project internships and make reasoned moral choices guided by the internal and external supervisors during the tenure. |  |  |  |  |  |  | ü |  |  |  |
| **CLO 4298.7** | Recognize the responsibilities deemed by the external and internal supervisors and understand the importance of supporting equality, diversity, and inclusion between peers. |  |  |  |  |  |  |  | ü |  |  |
| **CLO 4298.8** | Apply knowledge of engineering management principles and understand why project and change management may be required during practice school and project work. |  |  |  |  |  |  |  |  | ü |  |
| **CLO 4298.9** | Indicate the future direction of the project development and appreciate how it can be realized with collaborative, lifelong learning, and self-learning. |  |  |  |  |  |  |  |  |  | ü |

Declaration: Through this Industrial Training, I have accomplished above stated program articulation and IET learning outcomes.

**21**

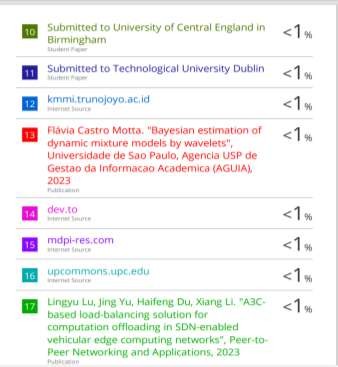
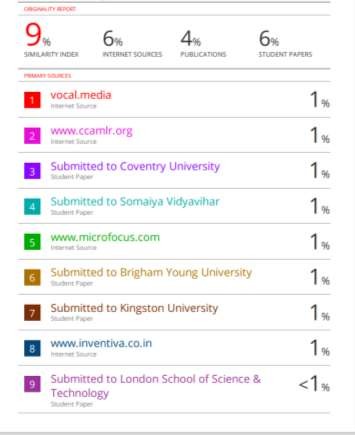
***REFERENCES***

*[1]*"Keras Documentation," [Online]. Available: <https://keras.io/guides/>

[2]“Transformers Documentation,” [Online]. Available: https://huggingface.co/docs/transformers/index

[3]“MNIST Dataset,”[Online]. Available: <https://www.tensorflow.org/datasets/catalog/mnist>

[4]”Self-Attention Layers”, [Online]. Available <https://armanasq.github.io/nlp/self-attention/>



**A close-up of a certificate

Description automatically generated**