

**Deep Learning**

***Project Title***

***Group 4***

*Ayan Choudhury, Andrew Dent, and Alicia Hernandez*

**School of Graduate Professional Studies**

Data Analytics

DAAN 570 – Deep Learning

*Fall, 2024*

**Document Control**

**Work carried out by:**

|  |  |  |
| --- | --- | --- |
| **Name** | **Email Address** | **Exhaustive list of Tasks** |
| Ayan Choudhury |  |  |
| Andrew Dant | Akd5358@psu.edu |  |
| Alicia Hernandez | Ash5978@psu.edu |  |

**Revision Sheet**

|  |  |
| --- | --- |
| **Date** | **Revision Description** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**TABLE OF CONTENTS**

[1 Introduction 3](#_Toc33129402)

[2 Problem Statement 3](#_Toc33129403)

[2.1 Challenges 3](#_Toc33129404)

[2.2 Related Works 3](#_Toc33129405)

[2.3 Importance 3](#_Toc33129406)

[3 Data Collection 3](#_Toc33129407)

[4 Data Preprocessing 3](#_Toc33129408)

[5 Methodology 4](#_Toc33129409)

[6 Results 4](#_Toc33129410)

[7 Discussion of Results 4](#_Toc33129411)

[8 References 5](#_Toc33129412)

# Introduction

[Enter Text Here.]

* Should state the background of your problem

# Problem Statement

[Enter Text Here.]

* Should state your problem statement

# Challenges

[Enter Text Here.]

* What are the challenges (business challenges, data and predictive analytics challenges?

## RELATED WORKS

[Enter Text Here.]

* Is the dataset already used in other projects?
* Introduce related projects and how your project and your solution differ from existing projects?
* Cite and add references to at least 5 scientific **papers** and **journals** related to your project and explain how their works differ from your contributions and findings. You can search the following databases through the **PSU libraries**: <https://libraries.psu.edu/databases>
  + ACM
  + IEEE
  + Springer
  + Elsevier
  + Web of Sciences
  + Google Scholar

## IMPORTANCE AND IMPACTS

[Enter Text Here.]

* Should state why you do think that your research question or the problem is important?
* Does it have any social/ economic/business/ scientific impact?

# Data Collection

[Enter Text Here.]

* Should mention the data source and the context in which the data is collected.
* Explain how your data is pertinent to your research question.
* Apply a Data Exploratory Analysis
* Give a snapshot of your data: feature description, types(categorical, continuous variables, ..)
* Preprocess features: missing values, skewness, type conversions, outliers, ..
* Visualize data (histograms, scatterplots, correlation, …)
* Interpret and explain your data
* …

# Data Preprocessing

[Enter Text Here. ]

* Apply descriptive statistics: means, std, min, max, range, ..
* Preprocess features: skewness, type conversions,
* Detect any inconsistency/ missing values in the data.
* You should mention whether scaling/ normalization / transformation of the features are required. If it is, then what you have done.
* Detect outliers, imbalanced or any inconsistent data, ….
* Check correlations and linearity., interpret your results ?
* Visualize and plot graphs/curves in support of your result..
* Interpret and explain your data
* Data Augmentation
* Data normalization and standardization, …
* …

# Methodology

[Enter Text Here. ]

* Should mention what methodology (ies) you have used in your project.
* Explain and justify the choice of your models
  + Why do you choose this(ese) model(s)?
  + Types of Neural Networks (RNN, CNN, sequential, functional API, …
    - Cost functions
    - Activation functions
    - Weights and biases initialization
* Set the strategy to validate your models
  + Pretrained models, transfer learning, …
  + Hold-out, Cross-validation, …)
  + Explain and present advantages / limits, ..
  + …
* How good is it (are they) in solving your problem?
* How do you compare your models to traditional machine learning?
* Mention about the assumption/constraints you consider while applying the method on your data.
* State what software/[Database] you have used

# Results and INterpretation

[Enter Text Here.]

* Interpret and explain the outcome of your models (coefficients, parameters, errors, etc...)
* Model performance and evaluations
  + Interpret and explain your results in terms
  + Accuracy, Losses, visual plotting, precision, recall, ROC etc.
  + Evaluation of Variance-Biases trade-off, how to improve them
  + Interpret your evaluation metrics for regression and/or classification
* You should provide the necessary tables that summarize your results and add charts.
* Figures, graphs/curves in support of your result.
* …

# Discussion of Results

[Enter Text Here]

* Discuss your results and its usefulness in solving your chosen problem.
* What will be the practical implication of your result (i.e. social/economic/commercial/scientific benefits) in society?
* What are the limitations (if any)?
* What would be the future work?
* What are the limits of your regressors or classifiers? What are your suggestions to improve them in the future?

# Your Feedback

[Enter Text Here]

* Feel free to share your feedback and thoughts about the final project terms and provide your positive feedback and suggestions to improve this project objectives.
* **…**

# References

[Enter Text Here].

* Add papers/publications/white papers citations.
* Add the links of the data/software you downloaded and when you visit them
* Add the relevant webpages/blogs you take help from.