Handwriting Digit Recognizer — Proposal

Capstone: The Art of Approximation
Ojas Chaturvedi, Ritwik Jayaraman, Saianshul Vishnubhaktula, Zaheen Jamil

Language

Python, a simple and popular language for machine learning and data science due to its extensive libraries and frameworks

Objective

To develop a custom machine learning model which would be able to determine what a digit is from an image of a handwritten single digit

Implementation

Overview of Steps:

1. Data Exploration and Visualization
2. Data Preprocessing
3. Feature Engineering
4. Model Building
5. Model Training and Testing
6. Model Evaluation and Deployment
7. Hyperparameter Tuning and Optimization
8. Website/App Development

Potential Libraries:

- Pandas: For data manipulation and analysis
- NumPy: For numerical computing and working with arrays
- Matplotlib: For data visualization
- Scikit-learn: For data mining and analysis
- TensorFlow: For deep learning and complex neural network modeling
- Flask/Django: For backend web development
- SQLAlchemy: For SQL databases and Object-Relational Mapping

Manual Work:

- Making algorithms for data preprocessing and feature engineering
- Building custom model
- Training and testing model
- Creating website/app that can use the model and store results for future training of model
- Documentation of all steps

Jobs

- Machine/Deep Learning Developers
  - Develops the machine learning model
  - Trains & tests the model
  - Makes the model usable in the website/app
- Data Analyst
  - Algorithm development for preprocessing and feature engineering
  - Will still contribute as a Machine Learning Developer
- GUI Developer
  - Makes the website/app and all of its functionality (UI)
  - Makes the model usable in the website/app
  - Will still contribute as a Machine Learning Developer