Peter Kamau Mwaura

AI for Software Engineering

Week 3: AI Tools Assignment

Part 1: Theoretical understanding

1. Short Answer Questions

Q1: Explain the primary differences between TensorFlow and PyTorch.

- **TensorFlow**: Developed by Google. Uses a **static computation graph** (define-and-run). Excellent for production deployment, with strong support for mobile and web (TensorFlow Lite, TensorFlow.js). Has a steeper learning curve but is very powerful and scalable.
- **PyTorch**: Developed by Facebook. Uses a **dynamic computation graph** (define-by-run), which is more intuitive and easier to debug, making it very popular in research. It feels more "Pythonic."
- **When to choose**: Choose **TensorFlow** for large-scale production systems and when you need robust deployment tools. Choose **PyTorch** for rapid prototyping, academic research, or if you prefer a more intuitive, Python-centric workflow.

Q2: Describe two use cases for Jupyter Notebooks in AI development.

- **Exploratory Data Analysis (EDA) and Prototyping**: Jupyter allows you to run code in isolated cells, making it ideal for incrementally loading data, visualizing distributions, handling missing values, and quickly testing small model ideas without rerunning the entire script.
- **Interactive Reporting and Education**: You can combine code, rich text, equations, and visualizations (like graphs and charts) in a single document. This is perfect for creating shareable reports that show both your process and your results, or for building tutorials where each step can be executed and explained.

Q3: How does spaCy enhance NLP tasks compared to basic Python string operations?

Basic string operations (e.g., `split()`, `find()`) are rigid and lack linguistic understanding. spaCy provides pre-trained statistical models that understand language context.

- **Example**: Finding people's names. With strings, you might look for capitalized words, which fails on "the johnsons" or includes "Paris" (a city). spaCy's NER model accurately identifies "PERSON", "ORG", "GPE" (locations) based on context and training.
- **Other Enhancements**: It handles tokenization, part-of-speech tagging, dependency parsing, and lemmatization with high accuracy and speed, which is nearly impossible to replicate with basic string ops.

2. Comparative Analysis: Scikit-learn vs. TensorFlow

Criteria	Scikit-learn	TensorFlow
Target Applications	**Classical ML:** Linear models, SVMs, decision trees, clustering, etc.	**Deep Learning:** Neural networks (CNNs, RNNs), large-scale complex data.
Ease of Use	**Very high.** Consistent API (`model.fit()`, `model.predict()`). Perfect for beginners in ML.	**Moderate to Steep.** Requires understanding of layers, optimizers, tensors. More concepts to grasp.
Community Support	**Excellent.** Huge community, extensive documentation, and many examples for classic algorithms.	**Massive.** Backed by Google, extremely active community, vast resources (blogs, courses, Stack Overflow).

Peter Kamau Mwaura

Part 2: Screenshots of the Model Outputs

Task 1: Classical ML with Scikit-learn

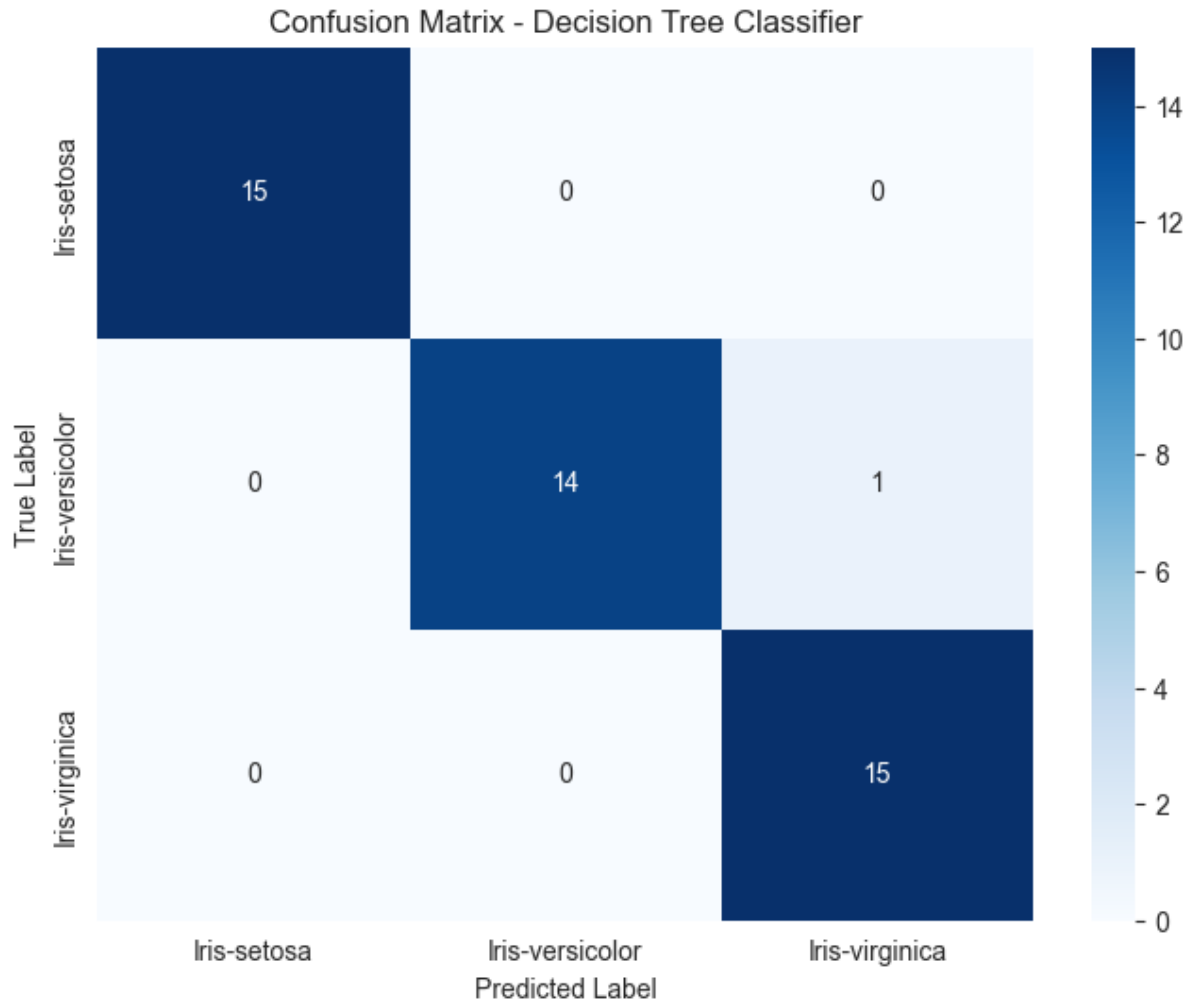

Figure 1:Confusion Matrix - Decision Tree Classifier

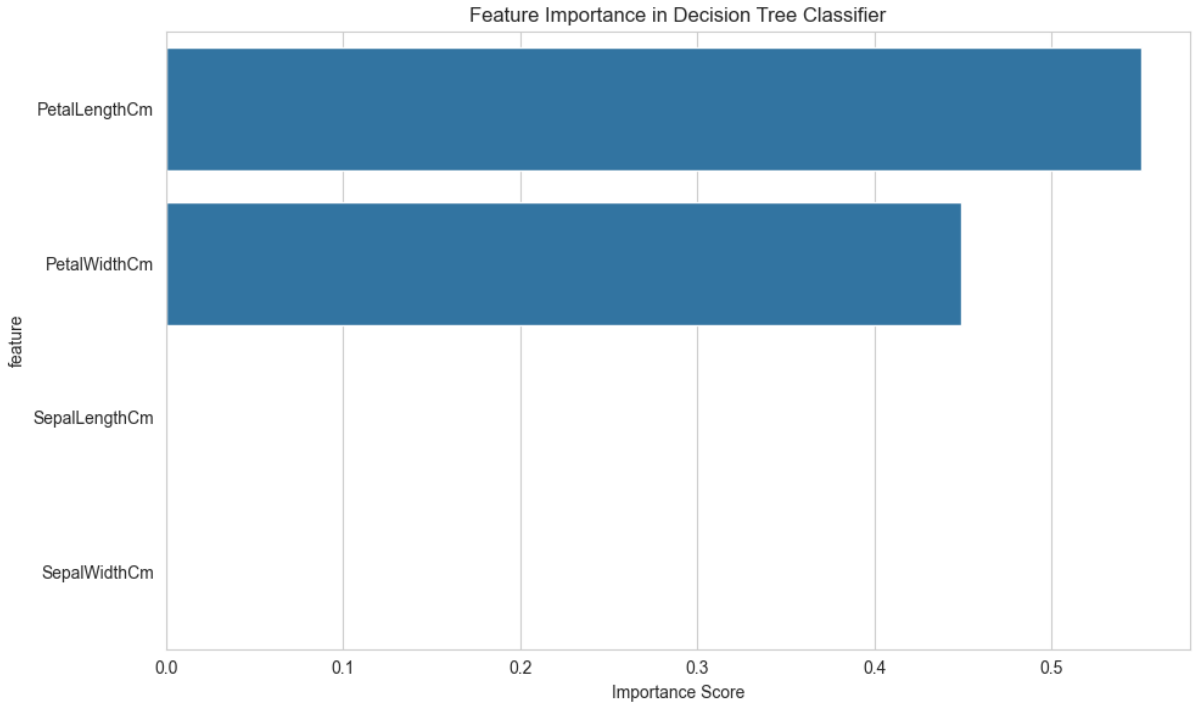

Figure 2: Feature Importance in Decision Tree Classifier

Task 2: Deep Learning with TensorFlow/PyTorch

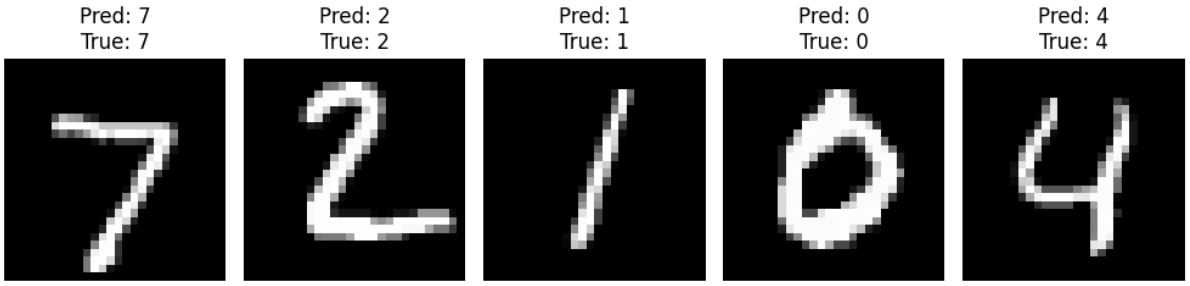

Figure 3: Visualizing sample predictions

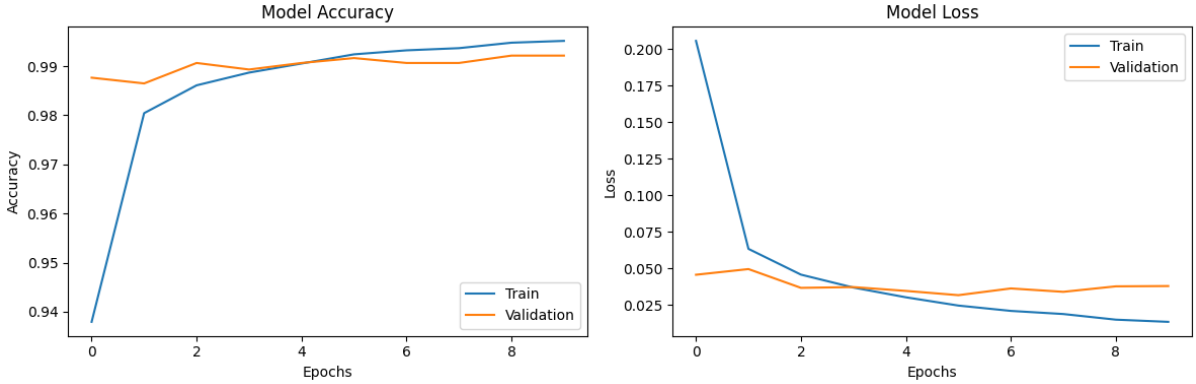

Figure 4: Plot Accuracy and Loss Curve

Task 3: NLP with spaCy

```
SUMMARY ANALYSIS
==================================================

Sentiment Distribution:
  POSITIVE: 8 reviews
  NEGATIVE: 1 reviews
  NEUTRAL: 1 reviews

Total Entities Found: 15

Detailed Results Table:
                                                                                                              review senti
ment  sentiment_score                                                        entities
              The Apple iPhone 15 has an amazing camera but the battery life is terrible. I expected better from Apple.  POSI
TIVE          0.2732                                     [iPhone (PRODUCT), Apple (ORG)]
Samsung Galaxy S23 is fantastic! The display is brilliant and the performance is smooth. Much better than Google Pixel.  POSI
TIVE          0.8908                                          [Samsung Galaxy S23 (ORG)]
                                  I love my new Kindle Paperwhite for reading books at night! The backlight is perfect.  POSI
TIVE          0.8478                         [Kindle (PRODUCT), Kindle Paperwhite (ORG)]
                                    The Sony headphones broke after just two weeks. Very disappointed with the quality.  NEGA
TIVE         -0.7346                                                       [Sony (ORG)]
                    Microsoft Surface Pro is a great device for work, but the price is too high compared to Apple iPad.  POSI
TIVE          0.3716 [iPad (PRODUCT), Microsoft Surface Pro (ORG), Apple iPad (ORG)]
                        Bought this Amazon Echo Dot and it's been working perfectly. Alexa understands all my commands.  POSI
TIVE          0.7096                                        [Echo (PRODUCT), Alexa (ORG)]
                          The battery on this Nintendo Switch doesn't last long. Otherwise, it's a good gaming console.  POSI
TIVE          0.4404                                           [Nintendo Switch (PRODUCT)]
                                                 Google Pixel camera is outstanding but the software has too many bugs.  POSI
TIVE          0.3612                                                                   []
                                           MacBook Air is lightweight and fast, perfect for students and professionals.  POSI
TIVE          0.5719                             [MacBook (PRODUCT), MacBook Air (ORG)]
                                            The Dell laptop overheats constantly and the customer service was unhelpful.   NEU
TRAL          0.0000                                                       [Dell (ORG)]

==================================================
ENTITY VISUALIZATION EXAMPLE
==================================================
Review: Apple iPhone and Samsung Galaxy are both great phones, but I prefer Google Pixel for its camera.

Entity visualization (run in Jupyter for better display):
Entities found:
  Apple iPhone - ORG
```

Figure 5: Summary Analysis

Peter Kamau Mwaura

Part 3: Ethical Considerations for MNIST and Amazon Reviews Models

Potential Biases in MNIST Model

The MNIST dataset consists of handwritten digits from 0-9, collected from various sources. Potential biases include:

- **Dataset Representation Bias**: The dataset may not represent all writing styles equally. For example, it might over-represent certain demographics (e.g., more samples from US postal workers) and under-represent others, leading to poorer performance on digits written by people from different cultural backgrounds or with disabilities.
- **Model Fairness Bias**: The CNN model might perform differently across subgroups. For instance, if trained primarily on clean, centered digits, it could struggle with noisy or off-center inputs, which might correlate with certain user groups.
- **Evaluation Bias**: Accuracy metrics might mask disparities; the model could achieve high overall accuracy but fail on specific digit classes or styles.

Mitigation Using TensorFlow Fairness Indicators

TensorFlow Fairness Indicators can help identify and mitigate biases:

- **Fairness Metrics**: Compute metrics like demographic parity, equal opportunity, and disparate impact across subgroups (e.g., by digit style or source).
- **Bias Detection**: Use tools to visualize performance disparities and identify biased predictions.
- **Model Adjustments**: Retrain or fine-tune the model with fairness constraints, such as adversarial debiasing or reweighting samples to balance representation.

Potential Biases in Amazon Reviews Model

The Amazon Reviews analysis uses spaCy for NER and VADER/TextBlob for sentiment. Biases include:

- **Sentiment Analysis Bias**: VADER and TextBlob are rule-based and may not handle sarcasm, context, or cultural nuances well. For example, positive words in negative contexts might be misclassified, and biases in training data (e.g., more positive reviews) could skew results.
- **NER Bias**: spaCy's en_core_web_sm model is trained on general English text, which may not recognize product-specific entities well or have biases towards certain brands (e.g., over-recognition of major tech companies like Apple vs. lesser-known brands).

- **Data Bias**: Reviews might not represent all demographics equally, leading to biased sentiment towards products popular in certain groups.

Mitigation Using spaCy's Rule-Based Systems

spaCy's rule-based systems can enhance fairness:

- **Custom Rules for NER**: Define patterns to recognize underrepresented products or brands, reducing bias in entity extraction.
- **Rule-Based Sentiment Adjustments**: Combine rule-based matching with ML to flag and correct biased sentiment predictions, e.g., rules for sarcasm detection.
- **Bias Audits**: Regularly audit the model on diverse datasets and adjust rules to ensure equitable performance across groups.

Overall, ethical AI requires ongoing monitoring, diverse data collection, and tools like these to promote fairness and reduce harm.