

## Group Members

Lynn Bitok

Adoh Baraza

Salome Wamaitha

## Q1: Differences between TensorFlow and PyTorch

Feature	TensorFlow	PyTorch
<b>Computation Graph</b>	Uses <i>static</i> (define-and-run) graphs (though TensorFlow 2 supports eager execution).	Uses <i>dynamic</i> (define-by-run) graphs, more intuitive for debugging.
<b>Ease of Use</b>	Steeper learning curve; more suited for large-scale deployment.	More Pythonic and user-friendly for research and prototyping.
<b>Deployment</b>	Excellent for production (via TensorFlow Serving, TF Lite, TensorFlow.js).	Deployment options exist (TorchServe, ONNX), but traditionally less extensive.
<b>Ecosystem</b>	Includes Keras API, TensorBoard, TF Hub, and TF Extended (TFX).	Integrated with PyTorch Lightning, TorchVision, and TorchText.

### When to choose:

- **TensorFlow** → Large-scale production systems, cross-platform deployment (mobile/web).
- **PyTorch** → Fast research prototyping, easier debugging, dynamic architectures.

---

## Q2: Use cases for Jupyter Notebooks in AI development

### 1. Interactive Experimentation & Visualization

- Ideal for testing ML models, plotting learning curves, and tuning hyperparameters interactively.
- Example: Visualizing CNN feature maps or PCA results.

### 2. Reproducible Research & Documentation

- Combines executable code, results, and markdown explanations in one document.
- Commonly used for presenting AI experiments or sharing tutorials.

---

## Q3: How spaCy enhances NLP tasks compared to basic Python string operations

- **Linguistic Awareness:**  
spaCy provides tokenization, POS tagging, named entity recognition, and dependency parsing — unlike simple `.split()` or regex operations.
- **Pretrained Models:**  
Includes efficient pretrained pipelines for multiple languages.

- **Performance & Integration:**

Built in Python for high speed and supports integration with ML frameworks (TensorFlow, PyTorch).

**In short:** spaCy transforms raw text into structured linguistic data, enabling advanced NLP tasks beyond simple string handling.

---

## 2. Comparative Analysis: Scikit-learn vs TensorFlow

Aspect	Scikit-learn	TensorFlow
<b>Target Applications</b>	Classical ML: regression, classification, clustering (e.g., SVM, Random Forest).	Deep learning: neural networks, CNNs, RNNs, transformers.
<b>Ease of Use for Beginners</b>	Very beginner-friendly; consistent APIs ( <code>fit()</code> , <code>predict()</code> ).	More complex; requires understanding of tensors, layers, and training loops.
<b>Community Support</b>	Mature, strong academic and enterprise adoption; excellent documentation.	Huge global community; widely used in production, research, and industry.

**Summary:**

- Use **Scikit-learn** for traditional ML tasks and quick prototyping.
- Use **TensorFlow** for deep learning or large-scale neural network models.

### Ethical Considerations

Model	Type of Bias	Description
<b>MNIST (Handwritten Digits)</b>	<i>Data Representation Bias</i>	The dataset contains primarily digits written in a particular style (Western handwriting). Models trained on it may perform poorly on digits written by people from different cultures or with disabilities.
	<i>Sampling Bias</i>	MNIST digits are clean and centered; the model might fail on real-world images with noise or rotation.
<b>Amazon Reviews Sentiment Model</b>	<i>Label Bias</i>	Sentiment labels may reflect annotators' subjective opinions.
	<i>Linguistic or Demographic Bias</i>	Reviews written in dialects or non-native English may be misclassified (e.g., "This movie was sick!" might be read as negative).
	<i>Imbalance Bias</i>	If positive reviews outnumber negative ones, the model learns to favor "positive" predictions.

### Bias Mitigation Using Tools

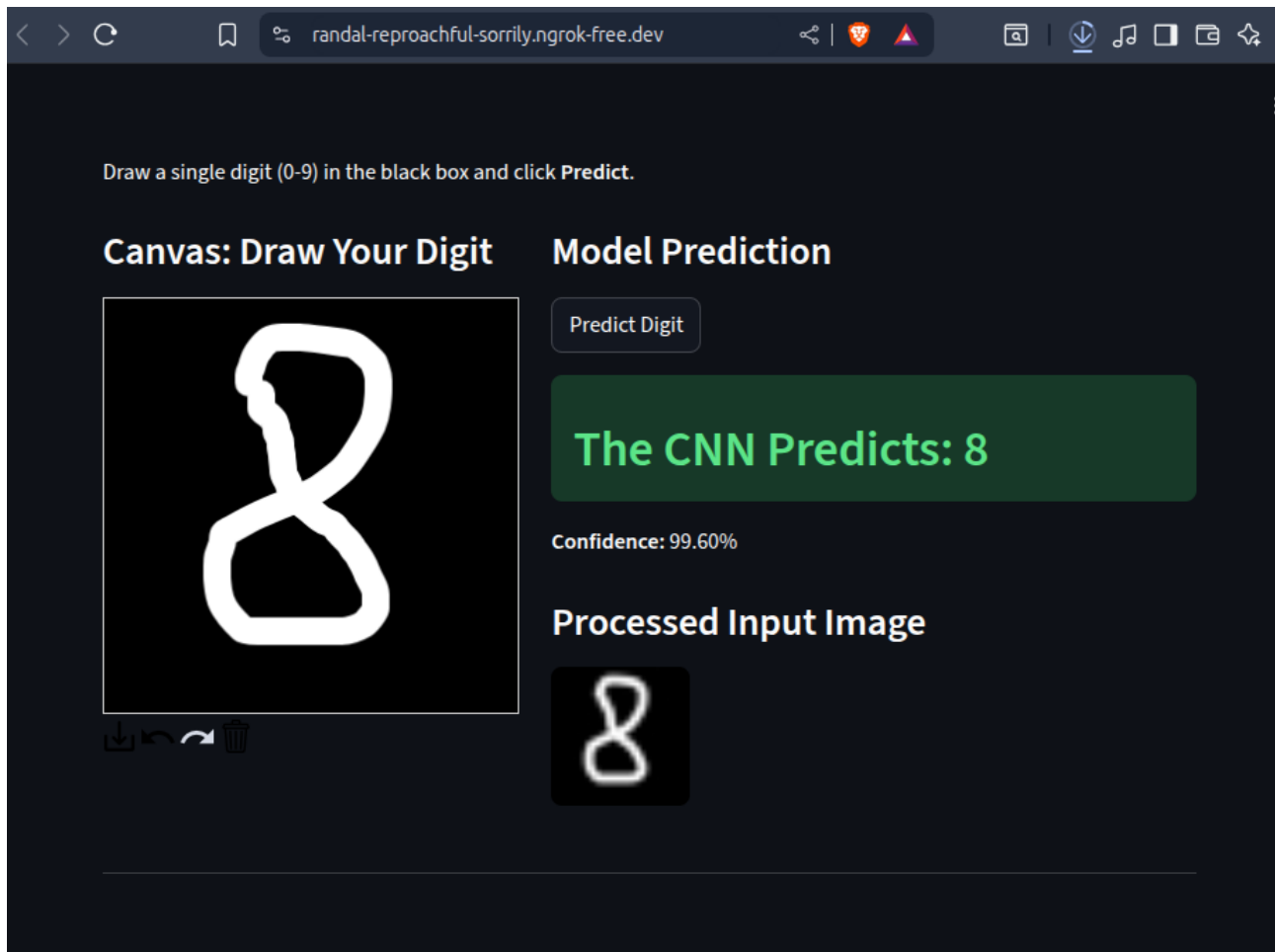
Tool	How It Helps
<b>TensorFlow Fairness Indicators</b>	Provides visual reports (like confusion matrices, false positive/negative rates) across <i>subgroups</i> (e.g., gender, dialect, region). → Detects whether accuracy is lower for specific demographic groups.

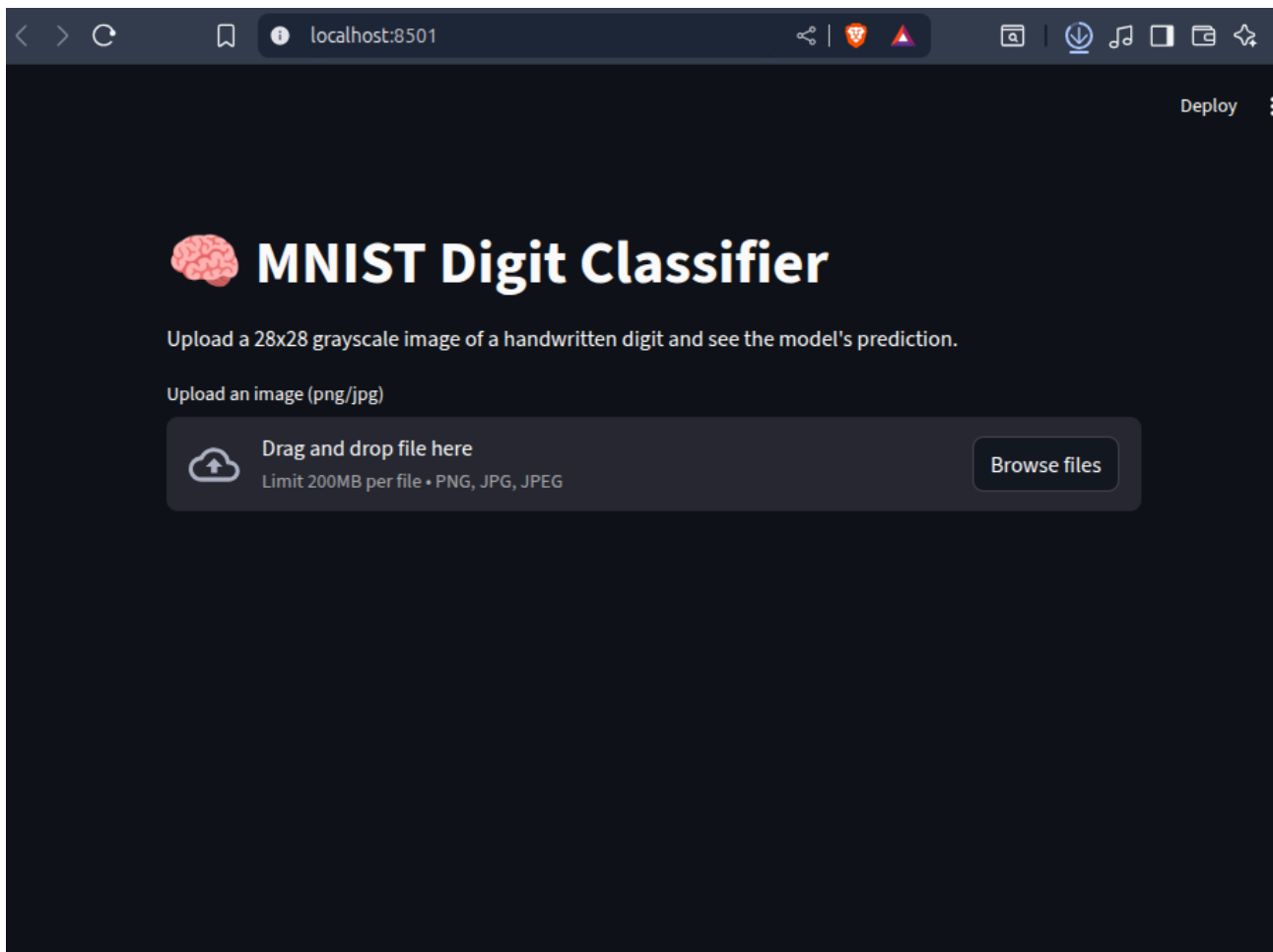
## Tool

### spaCy's Rule-Based Systems

## How It Helps

Allows you to enforce *linguistic normalization* before training (e.g., expanding slang, correcting contractions, or balancing named entity detection). →  
Reduces linguistic bias by ensuring consistent preprocessing.





Oct 19 19:28

TensorFlowDL Version control

TF Iris\_SKlearn.ipynb mnistCNN\_model.ipynb TextAnalysis\_Spacy.ipynb

Cell 1 is running Python 3.12.3: http://localhost:8888

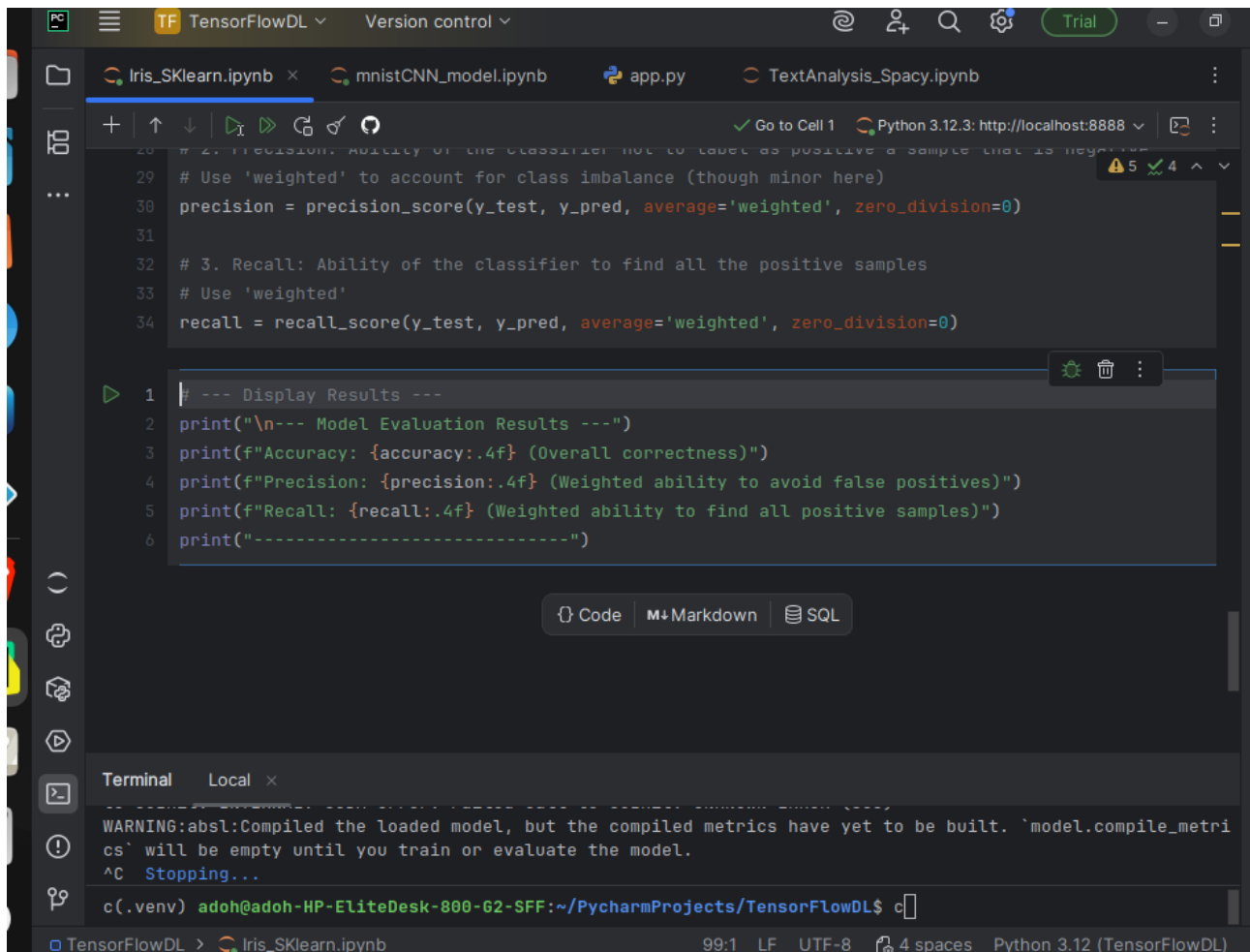
```
29     loss="sparse_categorical_crossentropy",
30     metrics=["accuracy"])
31
32 model.fit(x_train, y_train, epochs=5, batch_size=64, validation_split=0.1)
33
34 # 4. Evaluate
35 loss, acc = model.evaluate(x_test, y_test, verbose=0)
36 print(f"✓ Test accuracy: {acc * 100:.2f}%")
37
```

3m 24s

844/844 — 24s 29ms/step - accuracy: 0.9821 - loss: 0.0571 - val\_accuracy: 0.9878 - val\_loss: 0.0430  
Epoch 3/5  
844/844 — 25s 29ms/step - accuracy: 0.9875 - loss: 0.0398 - val\_accuracy: 0.9890 - val\_loss: 0.0404  
Epoch 4/5  
844/844 — 43s 31ms/step - accuracy: 0.9902 - loss: 0.0307 - val\_accuracy: 0.9888 - val\_loss: 0.0404  
Epoch 5/5  
842/844 — 0s 26ms/step - accuracy: 0.9935 - loss: 0.0210

```
1 # 5. Save model for Streamlit app
2 model.save("mnist_cnn_model.h5")
3 print("Model saved as mnist_cnn_model.h5")
```

TensorFlowDL > mnistCNN\_model.ipynb 2:1 LF UTF-8 4 spaces Python 3.12 (TensorFlowDL)



TF TensorFlowDLVersion control

Iris\_SKlearn.ipynbmnistCNN\_model.ipynbapp.pyTextAnalysis\_Spacy.ipynb

Go to Cell 1Python 3.12.3: http://localhost:8888

29

30 # Perform sentiment analysis using TextBlob

31 for review in reviews:

32 sentiment = TextBlob(review).sentiment.polarity

33 label = "Positive" if sentiment > 0 else "Negative" if sentiment < 0 else "Neutral"

34 print(f"Review: {review}")

35 print(f"Sentiment Score: {sentiment:.2f} → {label}\n")

✓ [3] 2s 23ms

Review: These Nike shoes are super comfortable and stylish.

Sentiment Score: 0.41 → Positive

Review: I had a bad experience with the Sony headphones; the sound was distorted.

Sentiment Score: -0.15 → Negative

Review: Amazon Echo is a great device, but Alexa sometimes misunderstands commands.

Sentiment Score: 0.80 → Positive

TerminalLocal × + ▾

Installing collected packages: en-core-web-sm  
Successfully installed en-core-web-sm-3.8.0  
✓ Download and installation successful  
You can now load the package via spacy.load('en\_core\_web\_sm')

(.venv) adoh@adoh-HP-EliteDesk-800-G2-SFF:~/PycharmProjects/TensorFlowDL\$

> TextAnalysis\_Spacy.ipUpdating skeletons...1:1LFUTF-84 spacesPython 3.12 (TensorFlowDL)