

Deep Learning Based Detection of Phishing Websites

Week 2 Report

David Toluwani Ehigie

February 27, 2026

1 Introduction

Week 2 focuses on developing deep learning models for phishing detection using raw URL text. Unlike Week 1, which relied on engineered features, this week explores how neural networks can automatically learn phishing patterns directly from URLs.

2 Objectives of Week 2

The main objectives were:

- Build a character-level LSTM model
- Improve the LSTM architecture
- Develop a CNN-based model for comparison
- Compare deep learning models with the Week 1 baseline

3 Dataset Representation

Only the URL column from the dataset was used in this week. URLs were treated as character sequences and converted into numerical format using character-level tokenization.

3.1 Tokenization

- Character-level tokenization was applied
- URLs were converted into sequences of integers
- Sequences were padded to a fixed length of 150 characters

4 Experiment 1: Baseline LSTM Model

A Long Short-Term Memory (LSTM) network was implemented to capture sequential patterns in URLs.

4.1 Initial LSTM Results

- Accuracy: 65.9%
- Precision: 92.4%
- Recall: 33.7%

The model showed high precision but low recall, indicating underfitting.

5 Experiment 2: Improved Bidirectional LSTM

The LSTM architecture was improved by:

- Increasing embedding size
- Increasing sequence length to 150
- Using Bidirectional LSTM
- Increasing training epochs to 10

5.1 Improved LSTM Results

- Accuracy: 87.3%
- Precision: 85.2%
- Recall: 89.8%

The improved architecture significantly enhanced performance.

6 Experiment 3: CNN Model for URL Classification

A Convolutional Neural Network (CNN) was implemented to capture local character patterns in URLs.

6.1 CNN Results

- Accuracy: 91.9%
- Precision: 91.9%
- Recall: 91.6%

The CNN outperformed the LSTM model, indicating that local URL patterns are highly informative for phishing detection.

7 Comparison with Week 1 Baseline

Model	Input Type	Accuracy
Random Forest	Engineered Features	96.9%
LSTM v1	URL Only	65.9%
Improved LSTM	URL Only	87.3%
CNN	URL Only	91.9%

8 Discussion

The experiments demonstrate that:

- Deep learning models can learn phishing patterns directly from URLs
- CNN performs better than LSTM for URL-based detection
- Engineered features still provide the highest accuracy

9 Conclusion and Next Steps

Week 2 successfully explored deep learning approaches for phishing detection. The CNN model achieved strong performance using only raw URLs.

Next Week: Development of a hybrid model combining engineered features with deep learning representations.