Using Modular Arithmetic Optimized Neural Networks to Crack Affine Cryptographic Schemes Efficiently

Vanja Stojanović, May 2025

UNIVERSITY OF LJUBLJANA
FRI
Faculty of Computer and Information Science

Motivation

- Affine ciphers (like classical ciphers) combine algebraic and statistical properties.
- Traditional neural cryptanalysis uses only statistical features.
- Can we do better by explicitly modeling modular arithmetic?

Affine Cipher Recap

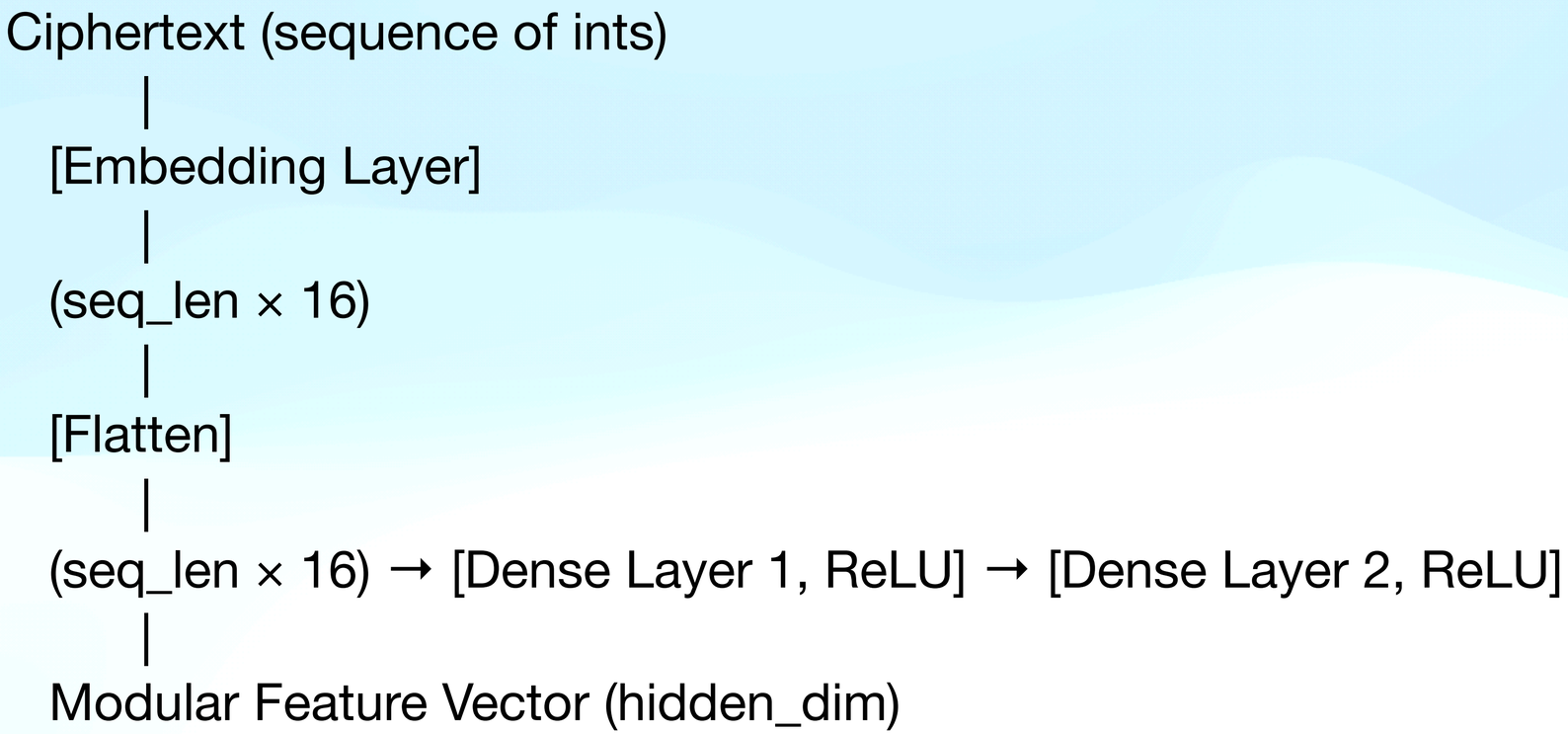
- Encryption is defined by $y = ax + b \mod 26$
- a is coprime to 26, b is any integer between 0-25
- The goal is to recover them through ciphertext only

Neural Network Architecture Overview

- Hybrid architecture: Modular branch + Statistical branch
- Both branches process the ciphertext in different ways
- Outputs are fused to predict the key

Modular Branch

- Learns modular arithmetic structure
- Steps:
  - Embedding: Each letter → 16D vector
  - Flatten sequence
  - Two dense layers (ReLU)
- Captures periodic and algebraic patterns


Ciphertext (sequence of ints)
[Embedding Layer]
(seq_len × 16)
[Flatten]
(seq_len × 16) → [Dense Layer 1, ReLU] → [Dense Layer 2, ReLU]
Modular Feature Vector (hidden_dim)

Statistical Brach

- Learns statistical properties of ciphertext
- Steps:
  - Compute letter frequency vector (26D)
  - Two dense layers (ReLU)
- Captures language statistics (e.g., common letters)

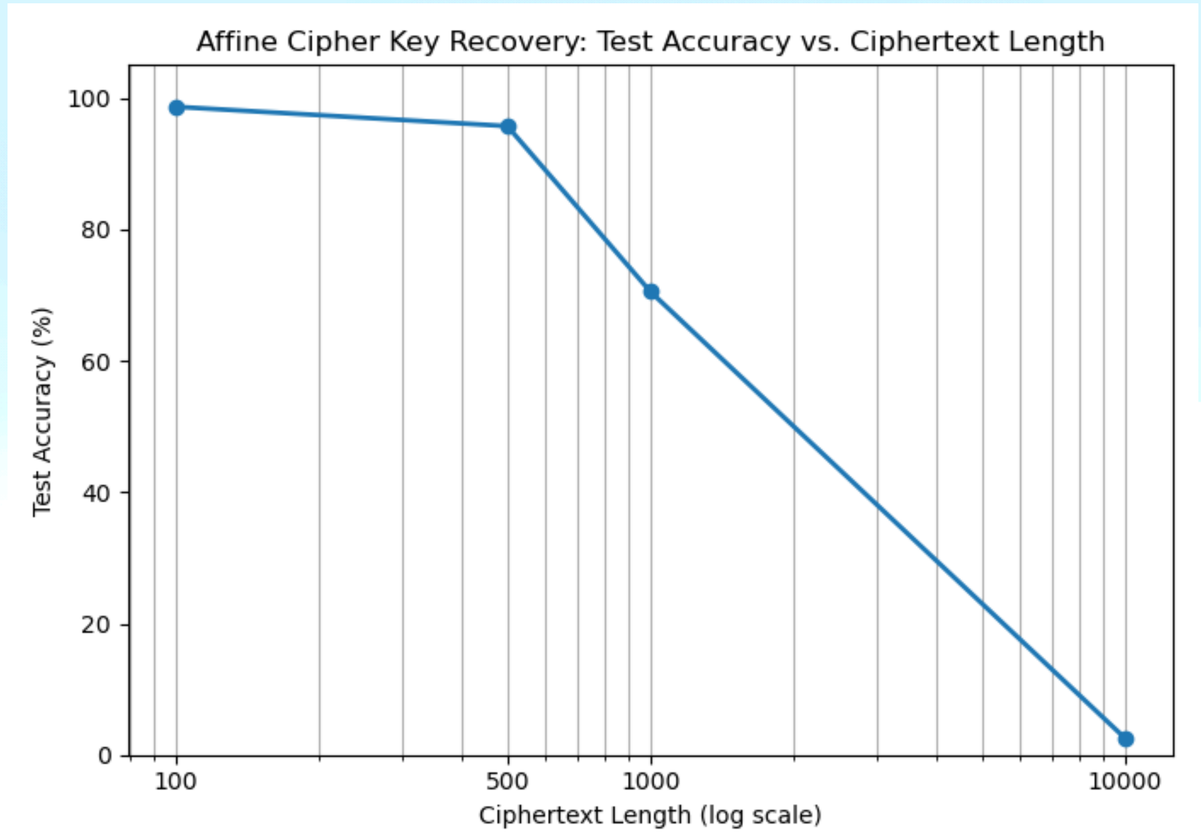
Fusion and Prediction

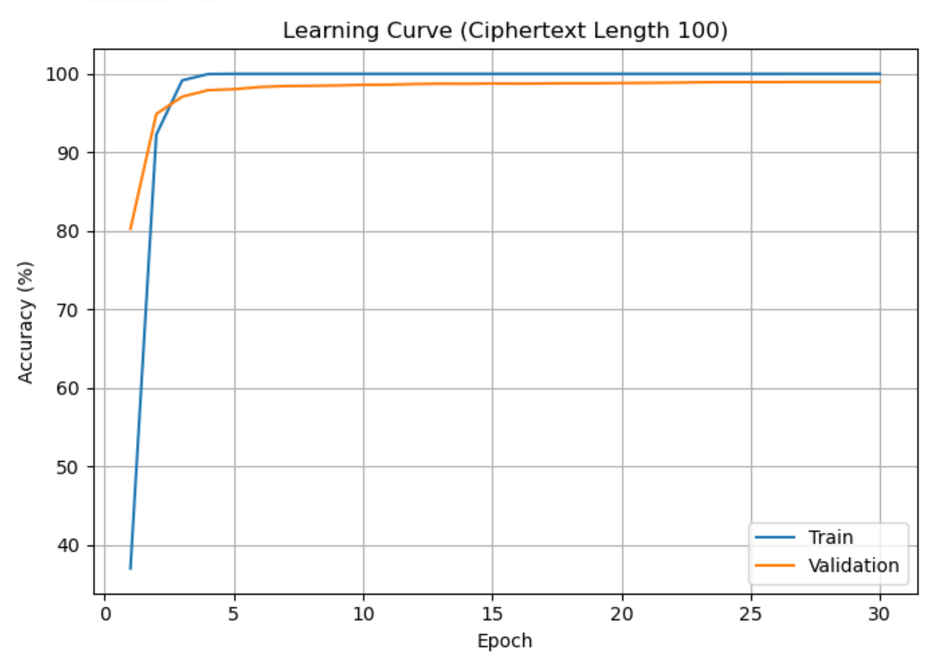
- Final dense layer predicts one of 312 possible keys
- Trained with cross-entropy loss

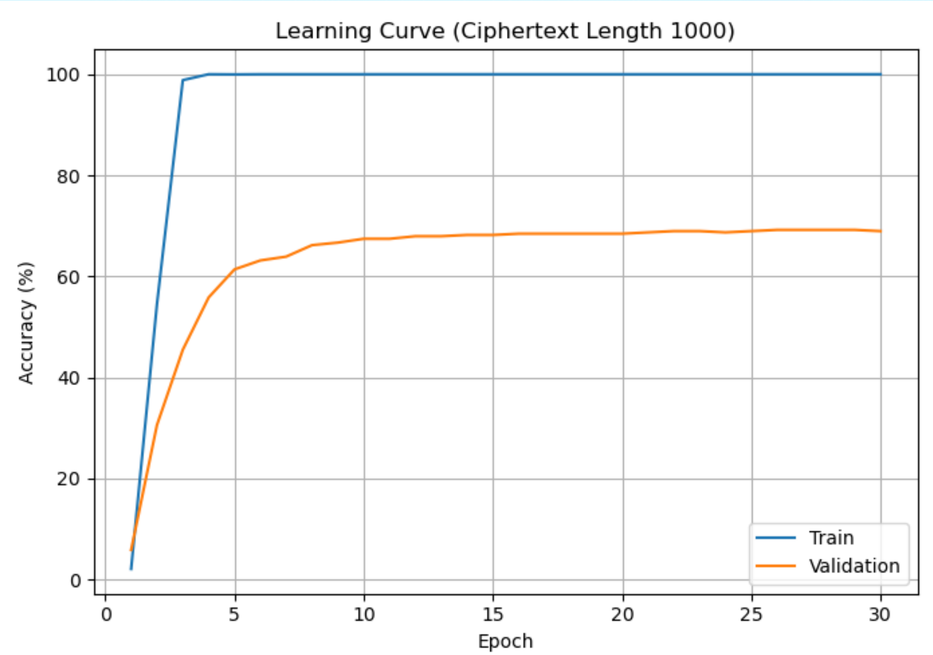
Training and Evaluation

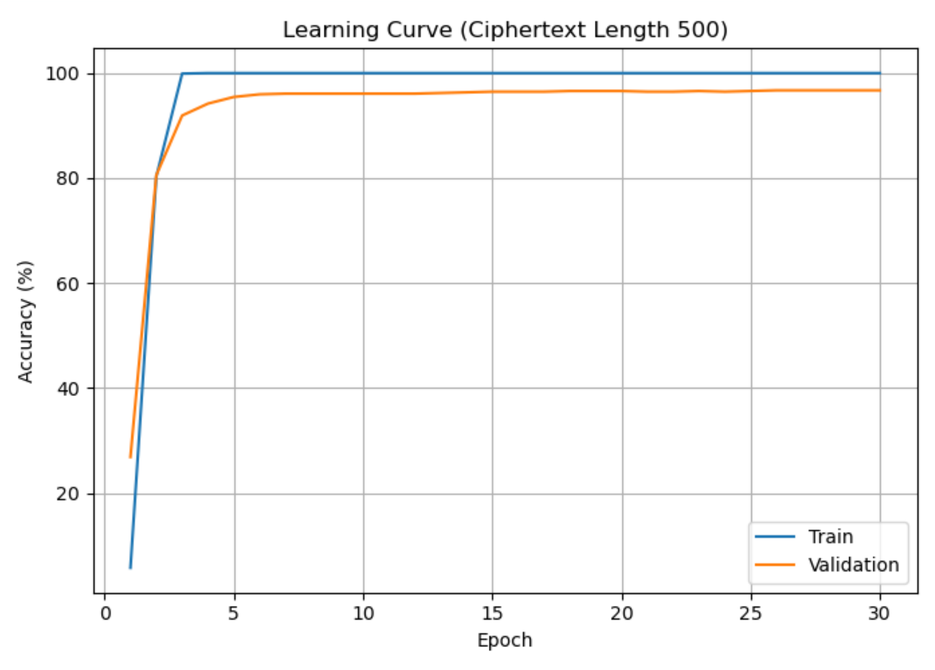
- Trained on labeled (ciphertext, key) pairs
- Evaluated on test set: accuracy of key recovery
- Hyperparameters:
  - Hidden size: 128
  - Batch size: 128
  - Learning rate: 0.0001
  - Epochs: 30

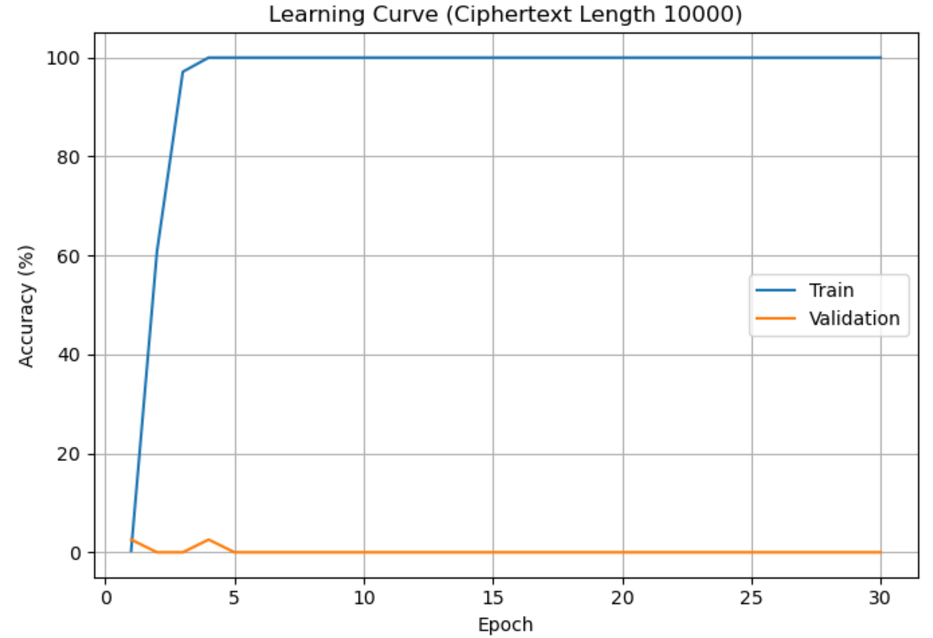
Test Accuracy vs. Ciphertext Length




Learning Curve (Ciphertext Length 100)
Accuracy (%)
Epoch
Train
Validation


Learning Curve (Ciphertext Length 1000)
Accuracy (%)
Epoch
Train
Validation


Learning Curve (Ciphertext Length 500)
Accuracy (%)
Epoch
Train
Validation


Learning Curve (Ciphertext Length 10000)
Accuracy (%)
Epoch
Train
Validation

Comparison with Prior Work

- Focardi & Luccio (2018): Used only statistical features
- Achieves higher or comparable accuracy, especially for affine ciphers

Cipher	Method (Focardi & Luccio)	Accuracy (short/moderate)	Accuracy (long)	Our Method (Affine)	Accuracy (short/moderate)	Accuracy (long)
Caesar	Statistical NN	100% (100–200 chars)	Not reported	N/A	N/A	N/A
Vigenère	Statistical NN	High	Not reported	N/A	N/A	N/A
Substitution	Statistical NN + search	~58% (200 words)	N/A	N/A	N/A	N/A
Affine	N/A	N/A	N/A	Hybrid NN	98% (100), 97% (500)	71% (1000), 2.5% (10000)