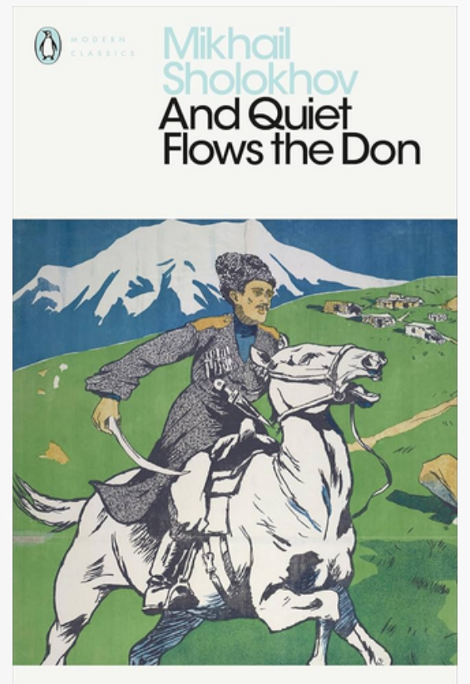
Analysis of Russian Texts Using Deep Impostor Method
25-2-R-12

BRAUDE
College of Engineering, Karmiel

By: Daniel Feldman and Arthur Cherniy
Supervisors: Dr. Renata Avros, Prof. Ze'ev Volkovich

The Problem

The authorship of "The Quiet Don", a novel written by Mikhail Sholokhov, a Nobel prize winning author, has been in debate since its publication in the late 1920s. Critics claim that parts of the novel may come from earlier Cossack writers. The debate was kept alive because of missing early drafts and the political climate of the Soviet Union, which made independent verification difficult.



The Need for the Development of the System

Despite decades of debate, the Sholokhov authorship question remains unresolved, largely because traditional arguments rely on subjective literary judgment. This creates a need for an independent,, data driven system that can evaluate authorship based on measurable text features rather than historical claims alone.

Requirements

- The system must store and manage large collections of literary texts, including disputed works, and impostor texts by authors of the same era.
- The system must support Russian text processing, tokenization, segmentation with minimal loss of stylistic information.
- The system must extract meaningful stylistic signals and represent texts in a form suitable for comparison.
- The system must identify texts that deviate from an author's established stylistic profile.
- The system must present results in a clear and interpretable form, enabling researchers to understand trends, outliers and confidence levels rather than just raw scores.

The Solution

To address the authorship problem, we developed a data driven authorship authentication system inspired by the Deep Impostor method proposed by Dr. Renata Avros and Prof. Zeev Volkovich. The system is designed to capture and compare an author's stylistic fingerprint across multiple works. For each impostor pair, a ConvBiLSTM model is trained to distinguish between text segments written by two impostors, while fine tuning a BERT model for text embeddings. Both are later used to classify the disputed works segments. Per iteration, each work is turned into a signal. After all iterations, a DTW distance matrix is computed for each impostor pair, and then Isolation Forest algorithm is applied to detect the anomalies. Afterwards, the signals are clustered based on their anomaly scores by using KMedoids algorithm. Based on the average of cluster assignments, each text receives a score between 1 (suspected fake) and 2 (authentic), with the threshold of 1.5 for authentic works .

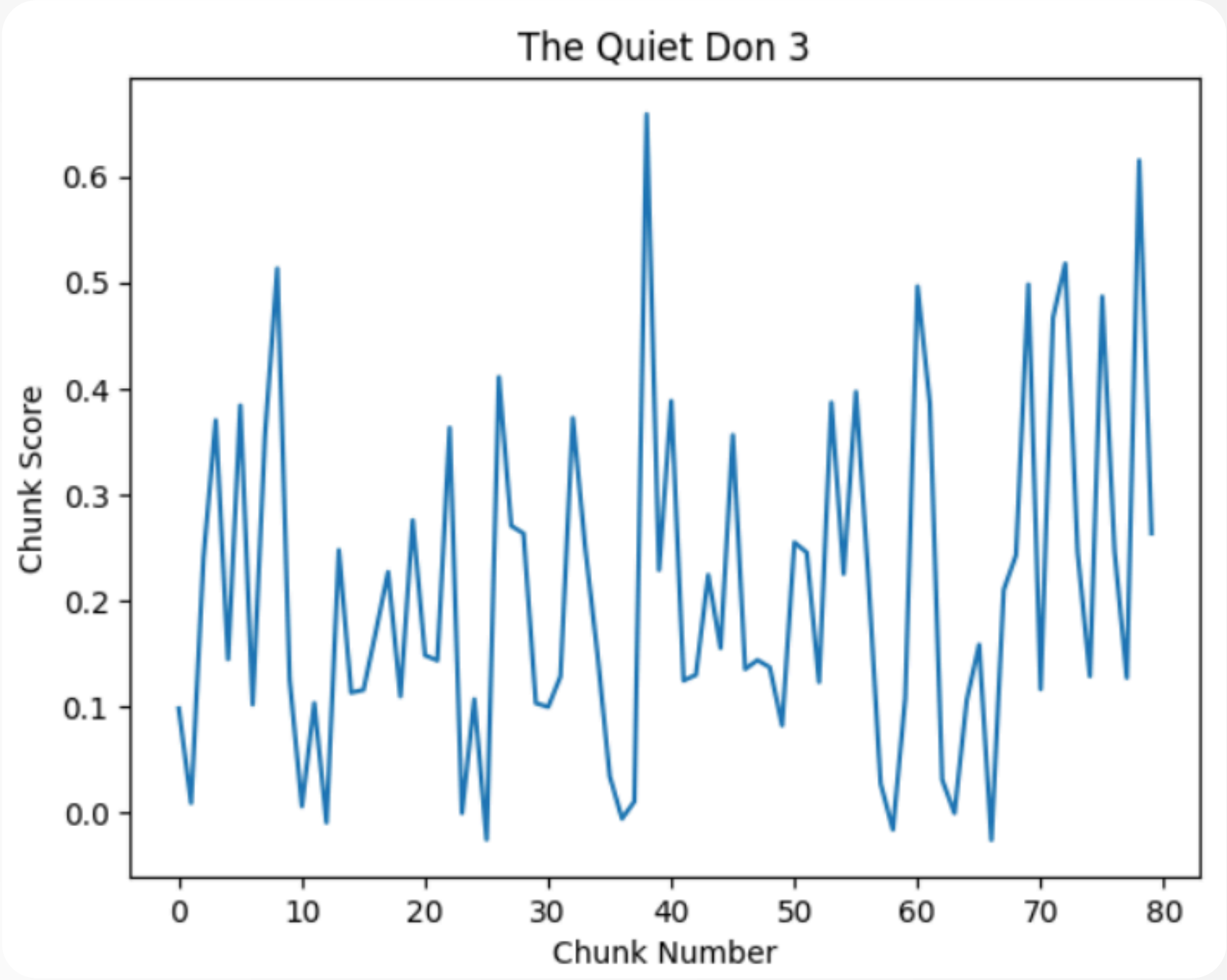


Figure 1. A signal representation of the book "The Quiet Don 3".

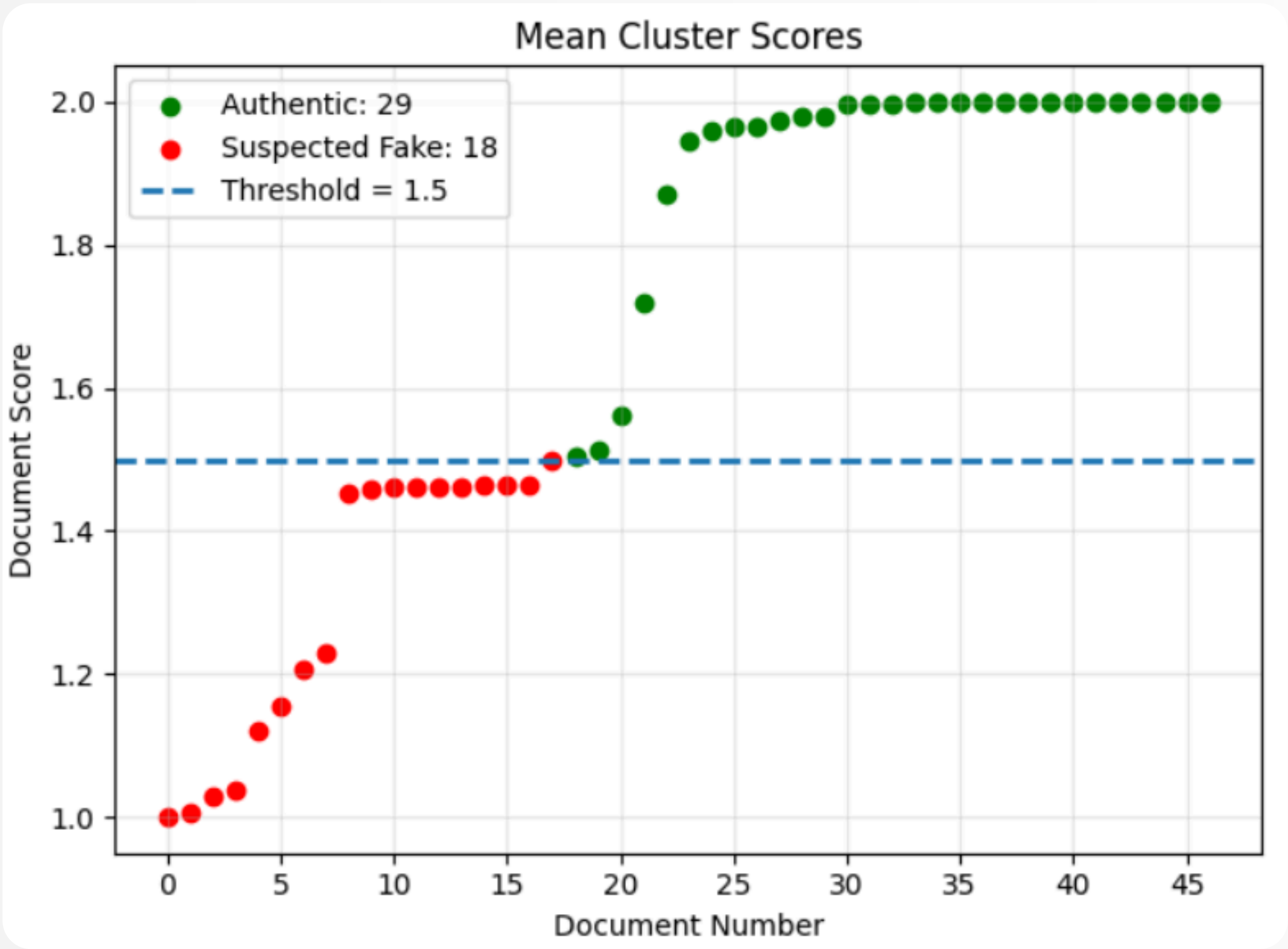


Figure 4. Mean cluster scores for each book, sorted from lowest to highest. **Red** indicates works suspected of plagiarism, while **green** indicates works considered authentic.

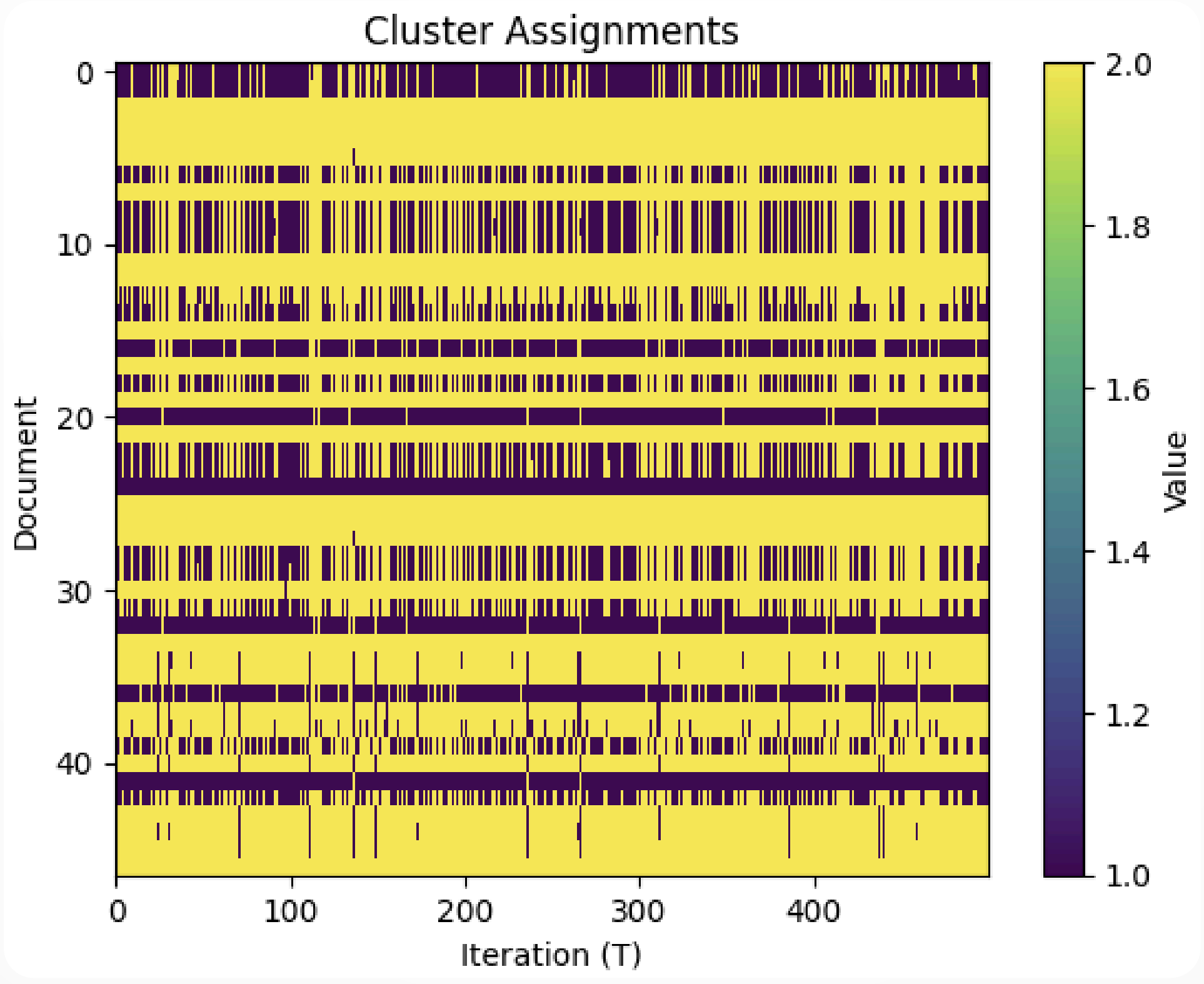


Figure 3. Cluster assignments per iteration, using the KMedoids algorithm with m=50 randomly selected features.

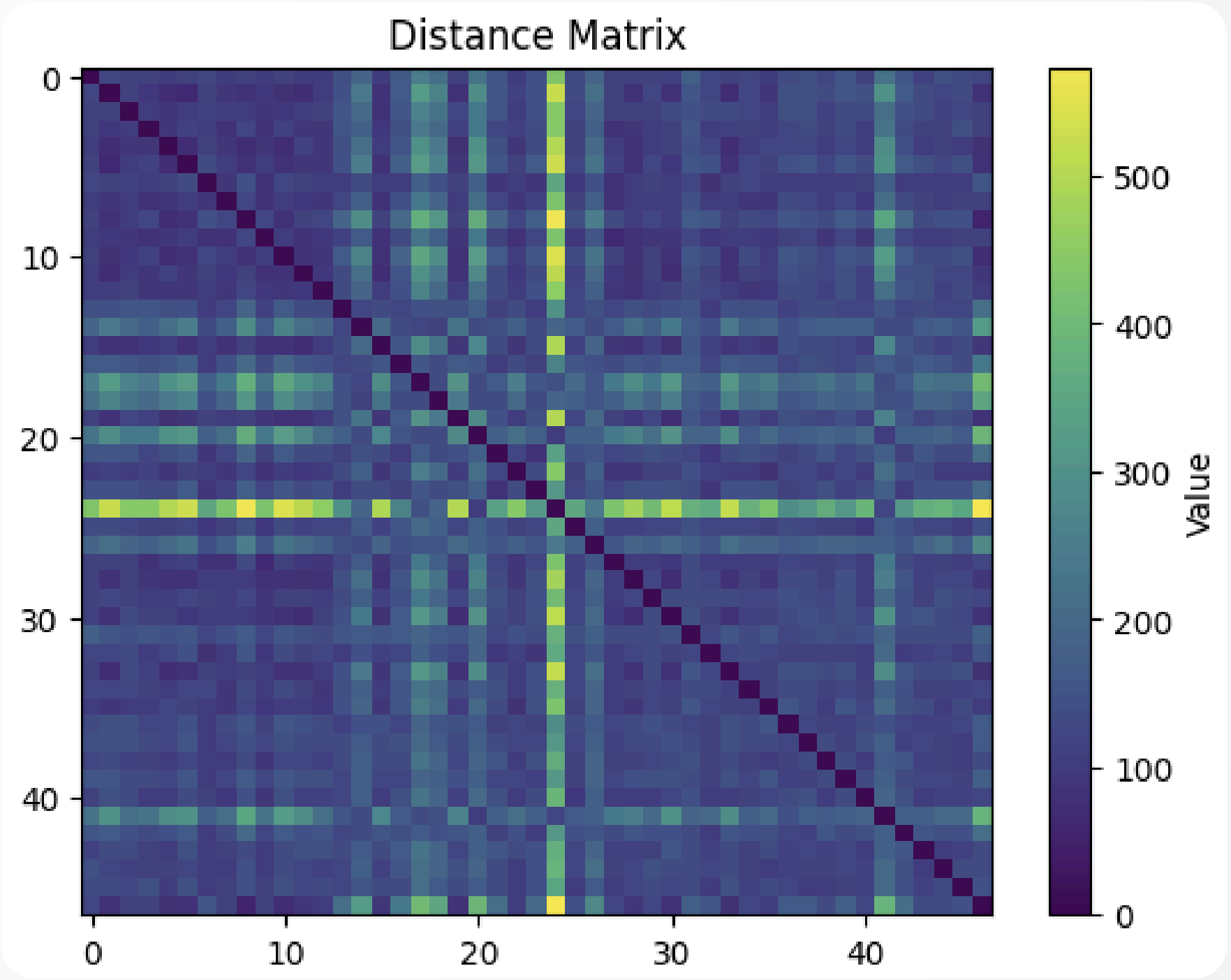


Figure 2. A pairwise distance matrix, used to save all the measured DTW distances from each signal to the rest of the signals, per iteration.

Evaluation Metrics

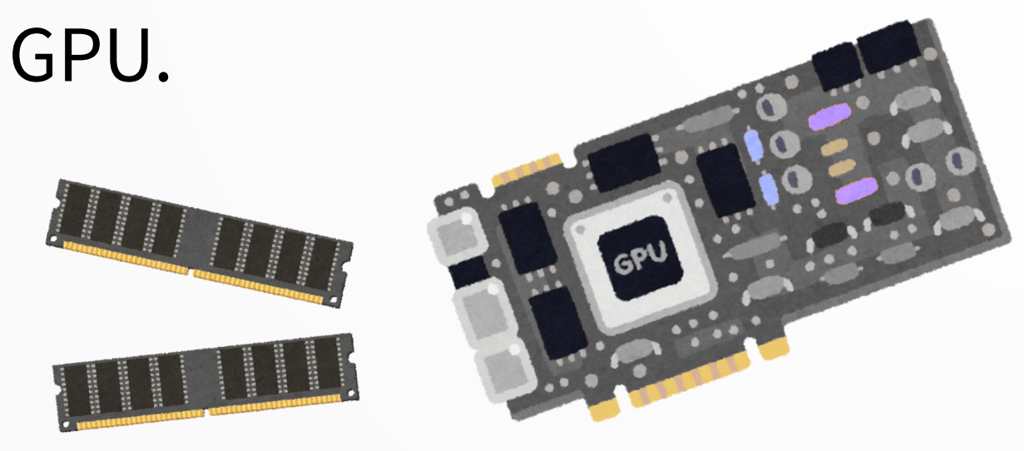
- An average accuracy of above 80% in prediction across all impostor pairs.
- Under 10 seconds for book prediction per impostor pair on average.
- Under 20 minutes for an impostor pair iteration.
- Results show at least partial alignment with conventional scholarly views on the authorship of the texts.

Meeting Metrics

- Reached an average accuracy of 87%.
- Book prediction per impostor pair takes 6.52 seconds on average.
- 12.34 minutes per impostor pair iteration on average.
- Results show partial alignment with the critics, indicating that "The Quiet Don" contains a different textual style.

Challenges

- CUDA out of memory issues - Fine tuning Bert models led to frequent Out Of Memory (OOM) issues.
- Prolonged Runtime - Authorship authentication is a very computation power and time demanding task, leading to very long runtimes.
- Downloading impostor corpus - manually downloading book files was a repetitive task.
- Device placement constraints - Bert models can only use input which is on the device they are on, whether it is the CPU or the GPU.
- NumPy and scikit-learn-extra incompatibility.



Actions Taken

- Addressed with improved memory management and upgrading to an A100 GPU.
- The program was broken down into 4 modules, each can be run separately on the saved output of the one before it in the chain.
- An automation script was written as a scraper service to download books from lib.ru.
- To address this issue we put extra effort into managing the data device location using `.to("cuda")` and `.cpu()`.
- A workaround was found where the colab session can be intentionally crashed in order to remove pre-installed incompatible dependencies and reinstall them.