Deep Learning to Detect Novel Behaviors in Traces

Quentin Fournier

Polytechnique Montréal
Laboratoire DORSAL

Novel Behaviors

- Novel behaviors are deviations from what has been previously observed:
  - Shift in normal behaviors: user patterns change over time.
  - New normal behaviors: new users or features.
  - Rare behaviors: unique queries.
  - Anomalous behaviors: latency, attacks, or bugs.

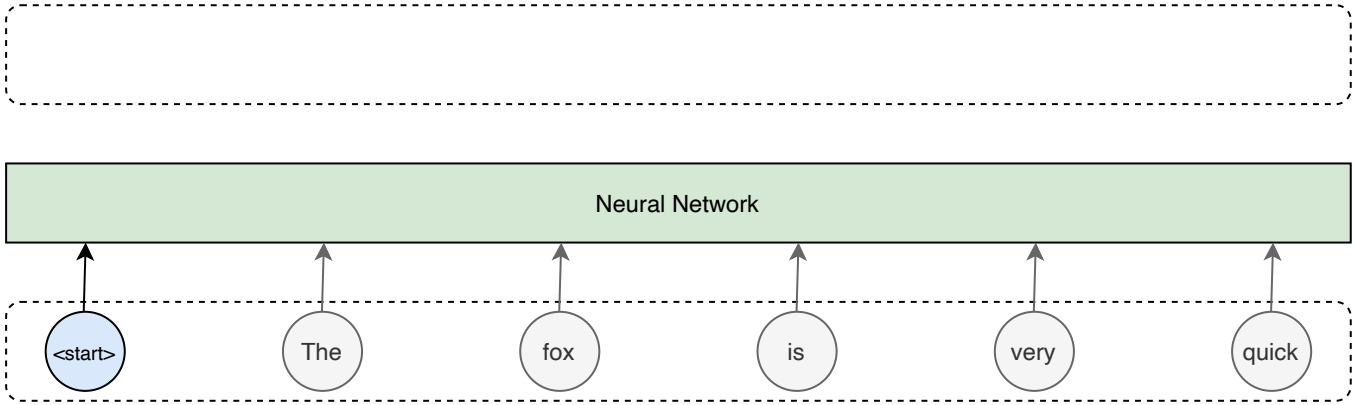
Approach - Dataset

- Sequences of system calls with their arguments:
  - Expose the system behavior.
  - Easy to collect large datasets.
  - Arguments enable more accurate predictions.

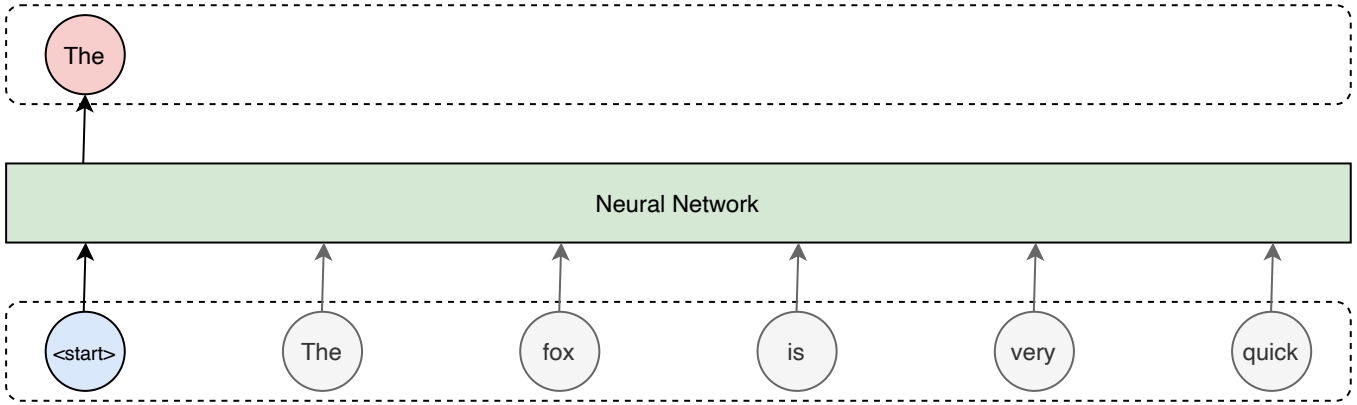
Approach - Objective

- Learn an unsupervised language model:
  - Labelling is time-consuming and error-prone.
  - Labels may change over time.
  - Predict the next system call knowing the previous ones.
  - Compute the likelihood of sequences.

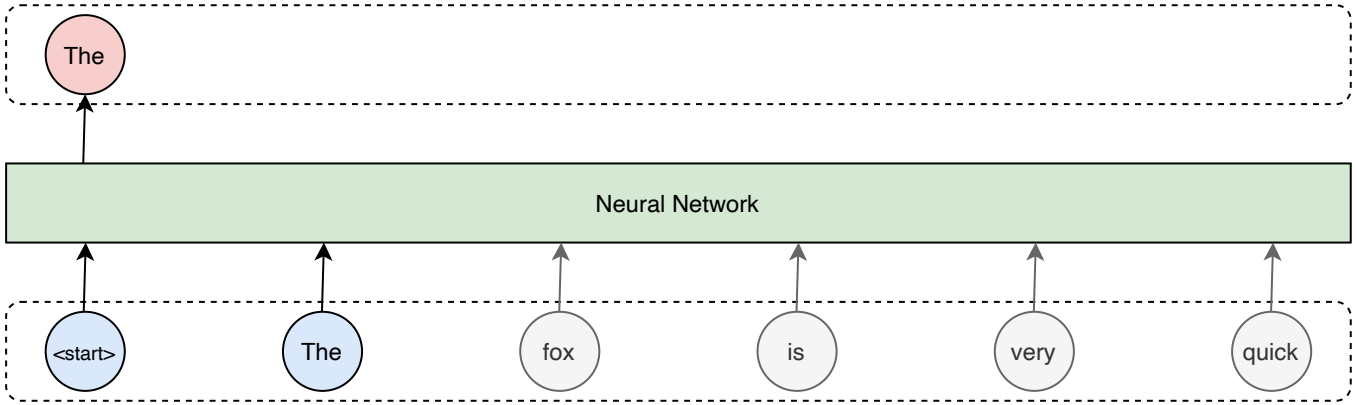
Approach - Objective



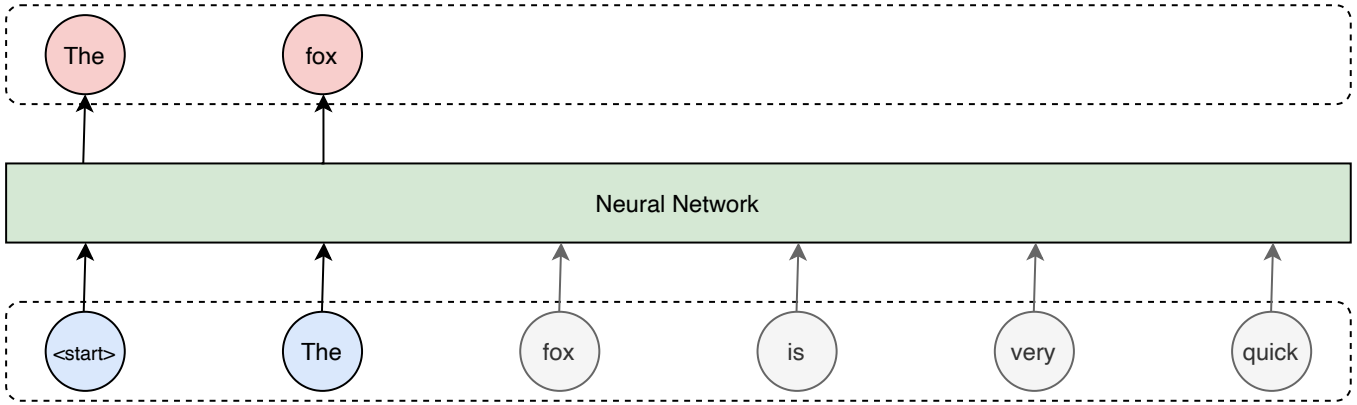
Approach - Objective



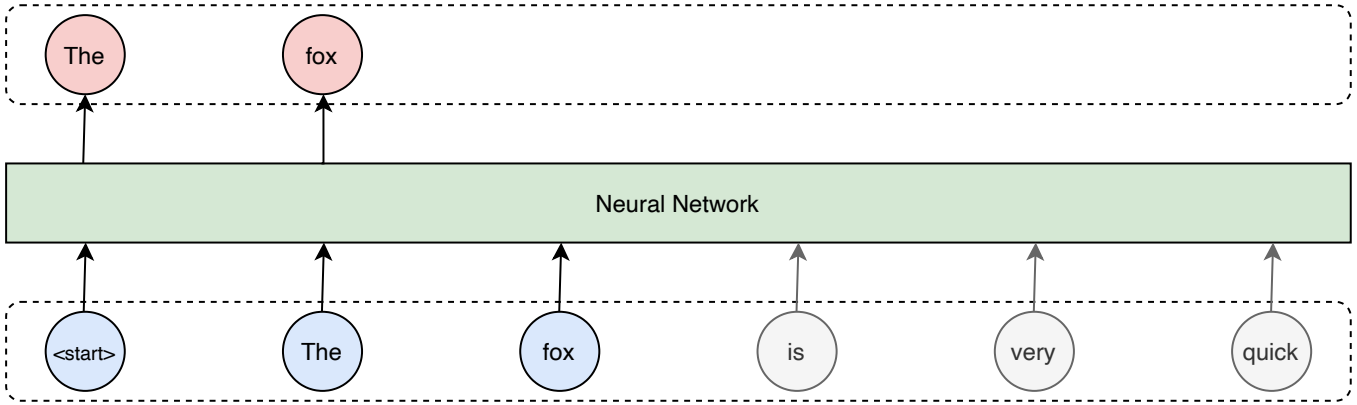
Approach - Objective



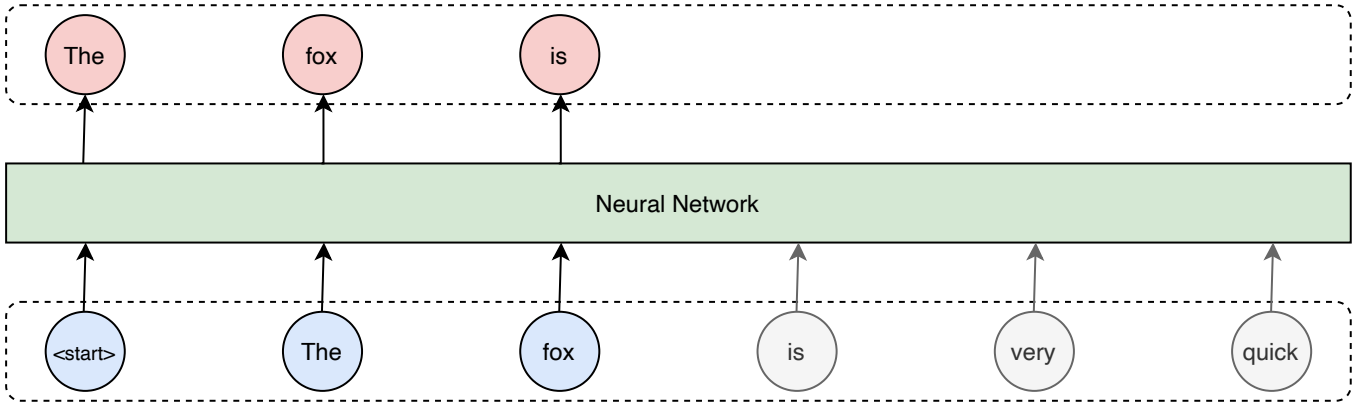
Approach - Objective



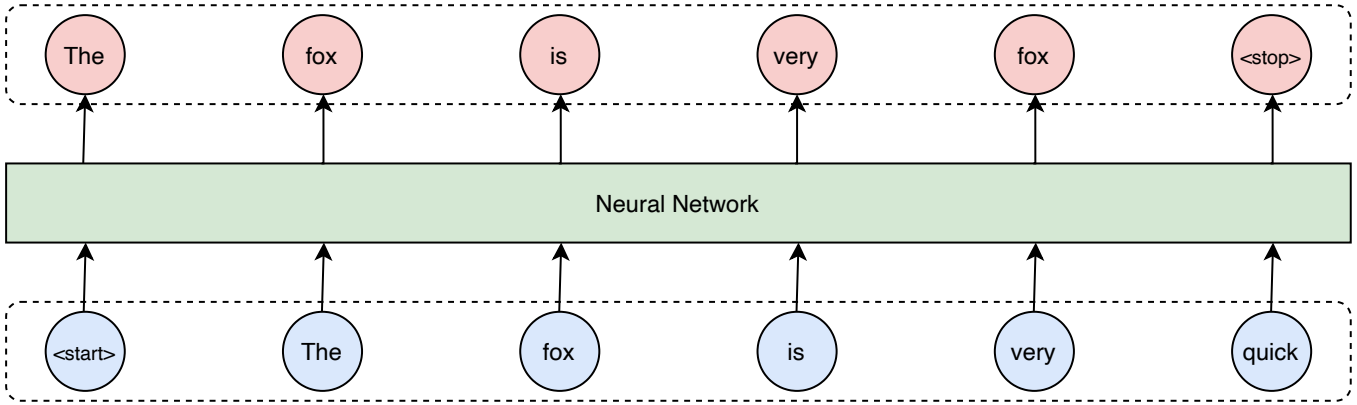
Approach - Objective



Approach - Objective



Approach - Objective



Perplexity

- The perplexity reflects how confused the model is about the input data.
- The higher the perplexity, the less likely the input data under the model.
- Set a threshold on the perplexity above which the data is considered unexpected and classified as novel behavior.

Approach - Models

- Finite state machine *n*-gram:
  - Conceptually simple and interpretable.
  - Unable to model complex and long-range interactions.
- Well known and proven neural network called LSTM:
  - Designed to process variable-size sequences.
  - Limited interaction range.

Approach - Models

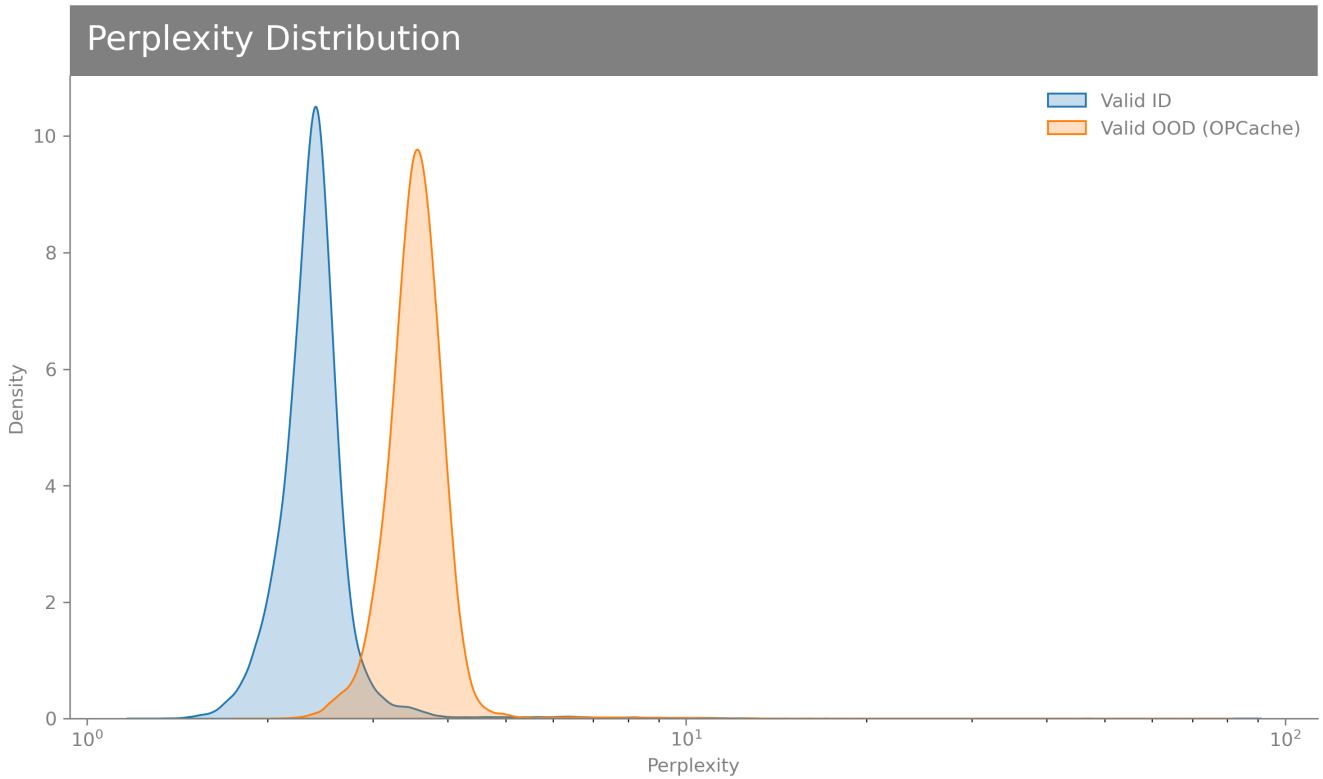
- State-of-the-art neural network called Transformer
  - Designed to model complex and long-range interactions.
  - Memory-intensive and sensitive to hyperparameters.
- An efficient Transformer called Longformer
  - Significantly less memory-intensive.
  - In theory, able to model complex and long-range interactions.

Results - Datasets

- 1 known behavior comprising 1,000,000+ web requests.
- 6 novel behaviors comprising 100,000+ web requests:each
  - Connection
  - CPU
  - IO
  - OPCache
  - Socket
  - SSL
- Simple and varied enough to evaluate the methodology.
- Too simple to represent large-scale real-world use cases.

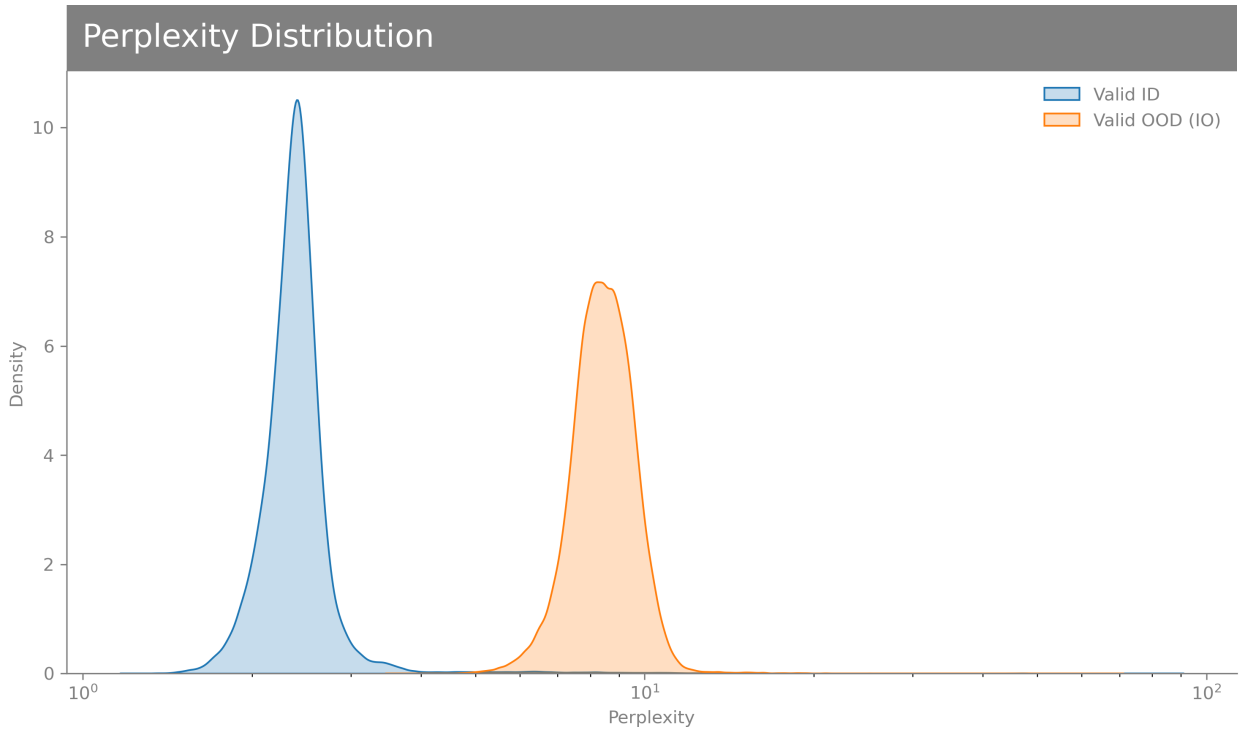
Results - LSTM - OPCache Disabled

- The novel behavior has a higher perplexity than the known behavior (small overlap):
- Accuracy of 95.56%.



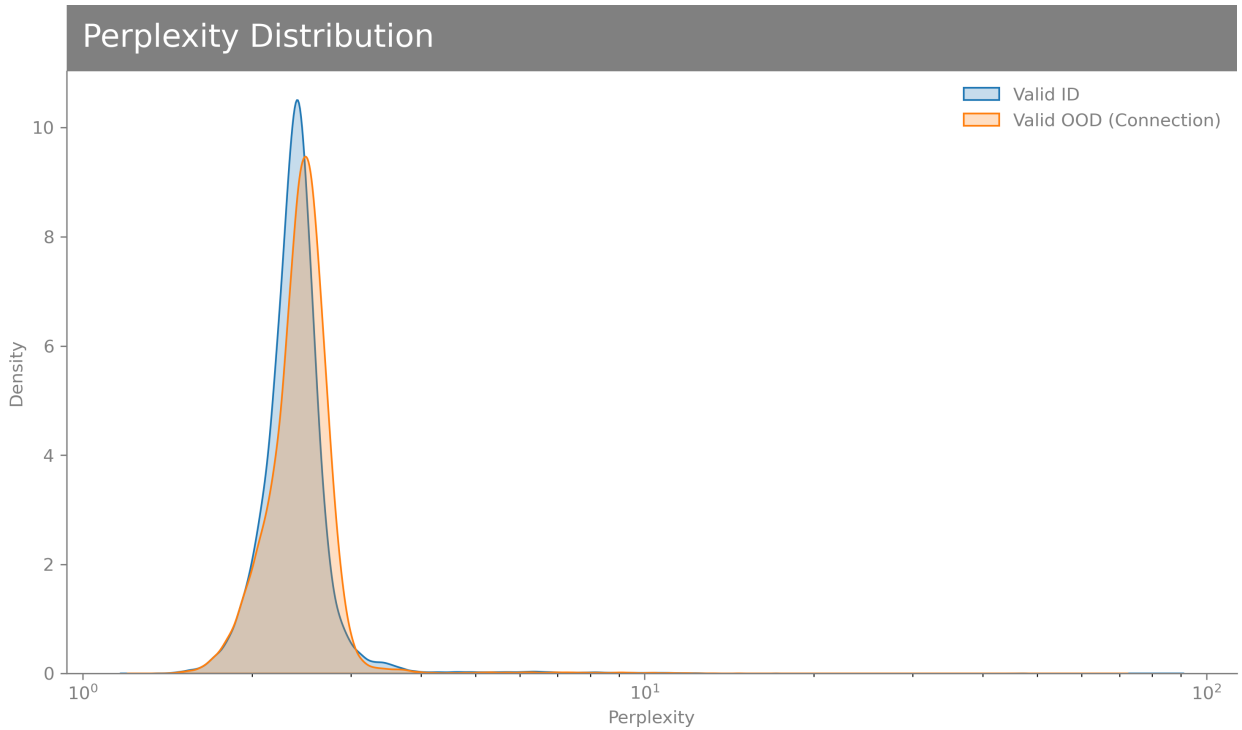
Results - LSTM - IO

- Best-case scenario, the novel behavior has a significantly higher perplexity (no overlap).
- Accuracy 99.48%:



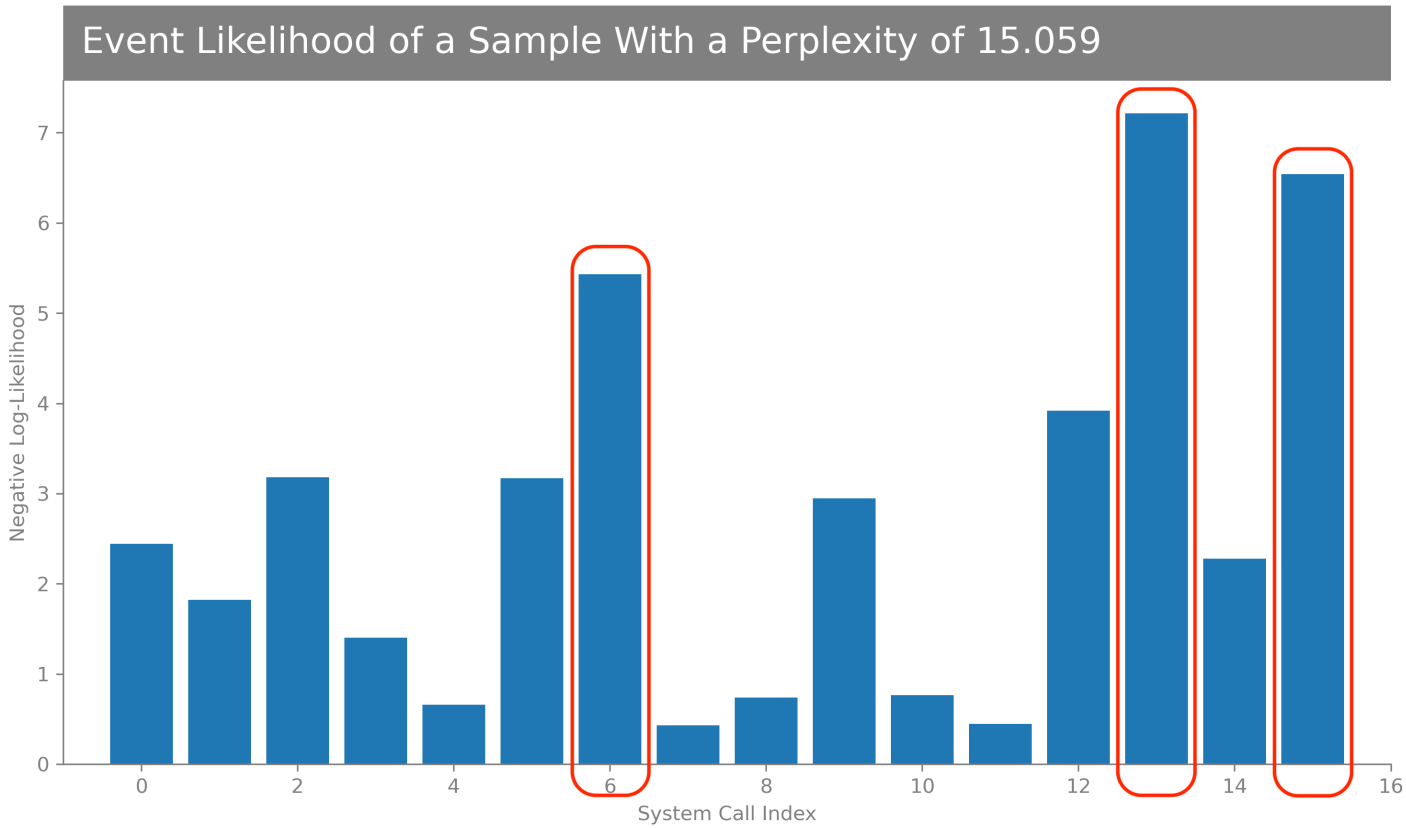
Results - LSTM - Connection

- Worst-case scenario, the novel behavior has the same perplexity as the known behaviors (large overlap).
- Accuracy 50.23% (random):



Results - Root Cause Visualization

- What system calls are responsible for the perplexity?



How to Continue the Project?

Limitations and Future Work

1. Most experiments were conducted on our own datasets due to the lack of publicly available large and modern datasets of kernel traces:
   - The datasets, the source code, the logs, and the trained models are open-source and are available on GitHub.
   - Experiment with your datasets.

Limitations and Future Work

2. No extensive hyperparameter search or ensemble due to time, cost, and environmental considerations:
    - The experiments are not resource-intensive and easy to reproduce yet achieve reasonable accuracy.

Limitations and Future Work

3. Three aspects of the proposed approaches must be improved before their deployment robustness, efficiency, explainability:
   a) Robustness: train an ensemble of networks.
   b) Efficiency: apply knowledge distillation.
   c) Explainability: visualize the attention and the conditional probability of the individual events.

Limitations and Future Work

4. Experiments were conducted with historical data, however, software and hardware continuously evolve:
   - One interesting avenue would be to consider lifelong learning.
   - The approach is compatible with online and real-time learning.