Model-Based AI Testing and Automation

Project Advisor: Jerry Gao

Computer Engineering Department

Tran, Chi (MS Software Engineering)
Kim, ILsoo (MS Software Engineering)
Tran, Hieu (MS Software Engineering)
Khajanchi, Deep (MS Software Engineering)

Introduction

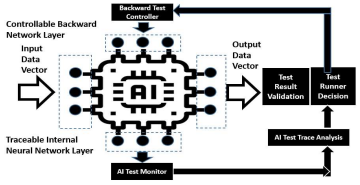
Artificial Intelligence (AI) is rapidly gaining more and more popularity in the modern world in different applications and businesses. Almost every smartphone currently has one or more AI-enabled personal assistants that are embedded in its operation system. Machine learning that applies neural networks is widely used in major business domains, such as social networks, robotics, and data science. By utilizing neural networks, people can manage time and the cost of operating big data-based applications as well as improve the accuracy of data analytics

In this project, the authors propose a model-based white-box testing methodology to investigate the relationship between given machine learning models and their model test coverage and accuracy coverage. The proposed solution includes visualizing neural networks' traverse inside each model, displaying linkages between nodes, and showing detailed information of each neural network layer. The test and accuracy coverage matrix also can be displayed to support users track, control, and enhance applications' quality.

Methodology

Proposed Areas of Study and Academic Contribution

The authors propose a solution that allows users to evaluate test coverage and accuracy coverage by accessing all layers inside neural networks to collect metrics, such as activation, weights, bias, loss, and accuracy.

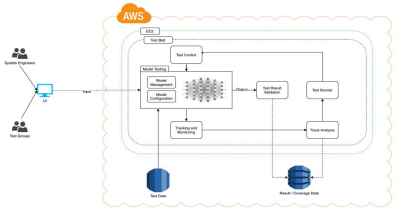


The proposed solution also provides visualization of neural networks by displaying all layers, all nodes in each layer, and linkages between nodes. Users can interact with layers by clicking on each one and going to the final node. The diagram below shows how the solution works to generate final reports.

Methodology

As presented in Figure 2, the input data vector is loaded into the models and will be processed through backward network layers. These layers will be controlled and traced under an AI test monitor to extract all necessary metrics to generate test analysis reports, test result validations, and test runner decisions. These tracings will be executed in loops correspondingly with feedback network layers. The proposed solution will be implemented on machine learning and deep learning projects that utilize LSTM and CNN models.

Project Architecture



- ***Model Testing***

The module performs predefined test scripts on the selected model and test data. When a request arrives, the Model Management submodule supplies the model, and the Model Configuration unit tweaks the model setting according to the request parameters. Initially, the models used in the project are neuron networks which have built-in feedback loops. Finally, the model testing unit pulls test data from an external NoSQL database and executes the scripts.

- ***Tracking and Monitoring***

This component collects trace data from trackers embedded in used models. Even though not every node in the network has a tracker installed, the tracked neurons should be distributed within the input layer, hidden layer and output layer.

- ***Trace Analysis***

The collected trace data are analyzed and aggregated in this module. The computed coverage data is written to an external SQL database.

- ***Test Runner***

The component decides the series of test scripts to run. These scripts can be predefined by system engineers or generated programmatically on-the-fly when needed.

- ***Test Control***

The test execution can be controlled any time with this unit. In addition to the basic functionalities such as scheduling / starting / pausing / stopping tests, the test control unit has to provide the ability to swap / insert scripts.

- ***Test Result Validation***

The output of the model is validated here. The results are written to the same SQL database that stores trace data. Various materialized views are created based on these two inserted data for faster querying.

Analysis and Results

The key features of the application

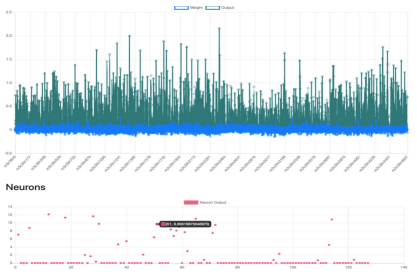
- Login
- Model management
- Test dataset management
- Neuron network visualization

Test results

- *User login*

- *Model management*

- *Neuron network visualization*



Summary/Conclusions

In the project, the authors have completed doing model management that allows users to deploy pre-trained models for checking accuracy, loss, weights, bias, and visualizing the fully-connected neuron networks between layers.

The project supports users to analyze the training results and improve quality of model training. Moreover, users can display all misclassified images with heatmaps and activations.

Key References

[1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, New Jersey, 1995.

[2] Imre J. Rudas, János Fodor, "Intelligent Systems", *Int. J. of Computers, Communications & Control*, ISSN 1841-9836, E-ISSN 1841-9844 Vol. III (2008), Suppl. Issue: Proceedings of ICCCC 2008, pp. 132-138.

[3] R.C. Eberhart and Y. Shui, Computational Intelligence – Concepts to Implementations, *Elsevier*, 2007.

[4] Xingjian Shi Zhourong Chen Hao Wang Dit-Yan Yeung, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." *Advances in neural information processing systems*. 2015.

[5] J. Gao, C. Tao, D. Jie and S. Lu, "Invited Paper: What is AI Software Testing? and Why," *2019 IEEE International Conference on Service-Oriented System Engineering (SOSE)*, San Francisco East Bay, CA, USA, 2019, pp. 27-2709.

[6] C. Tao, J. Gao and T. Wang, "Testing and Quality Validation for AI Software–Perspectives, Issues, and Practices," in *IEEE Access*, vol. 7, pp. 120164-120175, 2019.

Acknowledgements

The authors are deeply indebted to our Advisor, Prof. Jerry Gao for his invaluable comments and assistance in the Preparation and development of this project.

We would also thank Prof. Dan Harkey very much for his support and guidance with the project.

We would love to send our gratefulness to our family and friends for their moral support and encouragement.