MSCPES 2020 WORKSHOP
APRIL 21, 2020

Online Testbed for Evaluating Vulnerability of Deep Learning Based Power Grid Load Forecasters

Authors: Himanshu Neema, Peter Volgyesi, Xenofon Koutsoukos, Thomas Roth, Cuong Nguyen

Presented By:
Dr. Himanshu Neema
Research Assistant Professor
Vanderbilt University
Email:
himanshu.neema@vanderbilt.edu

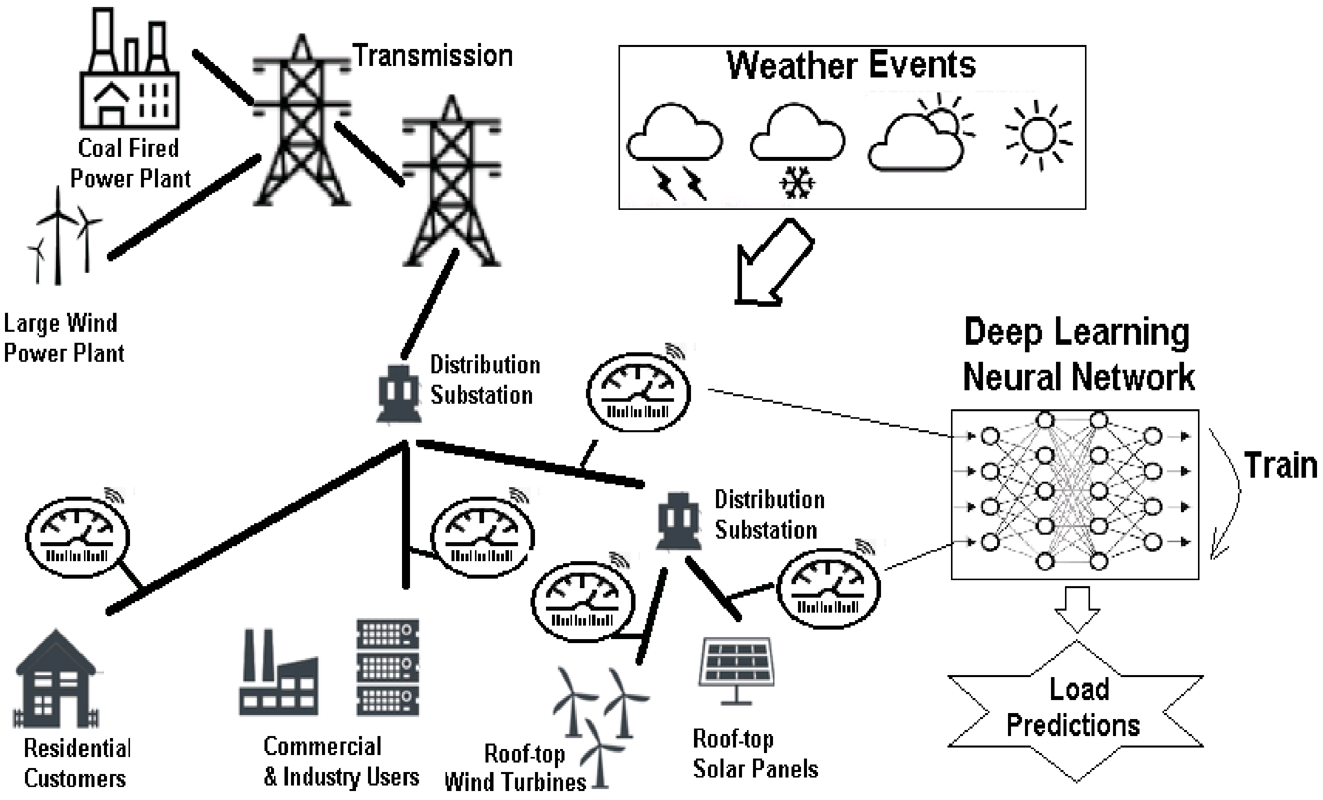
Acknowledgements:

- National Security Agency (NSA)
- National Science Foundation (NSF)
- National Institute of Standards and Technology (NIST)

VANDERBILT UNIVERSITY

NIST
National Institute of Standards and Technology

Machine Learning in Power Grid Load-Forecasting



- Distributed Energy Resources (DER) integration makes grid controls highly dynamic and distributed
- Prosumers = Producers + Consumers
- Dynamic power pricing adds to complexity
- Traditional load forecasting becomes highly challenging
- Deep-learning based predictors using smart meter data is more manageable

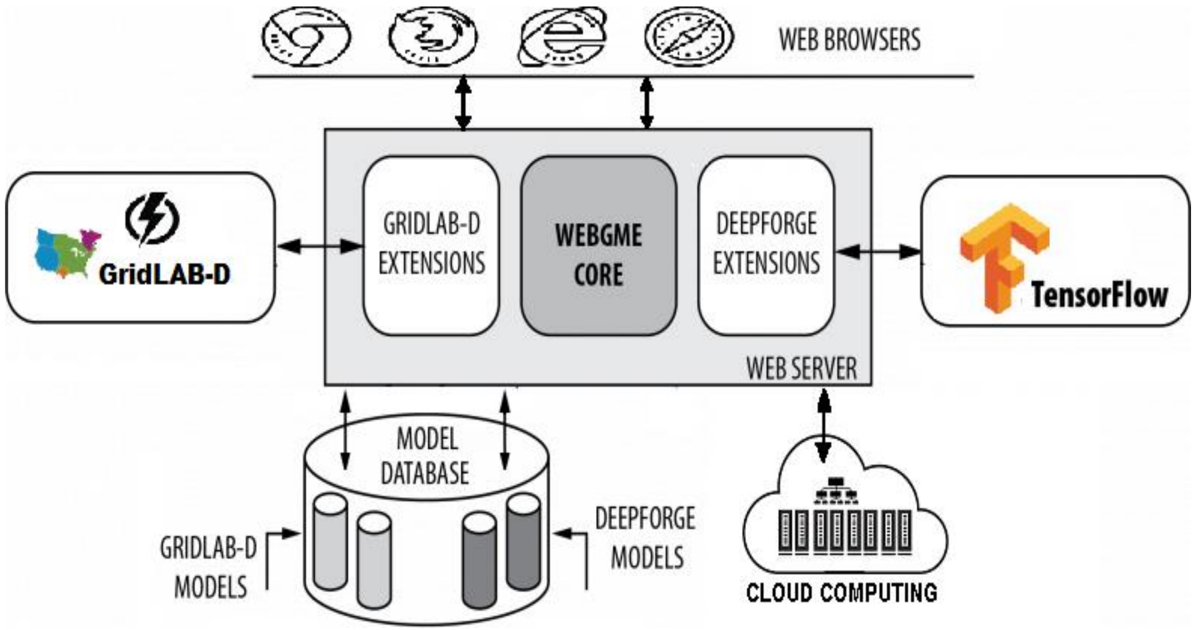
PROBLEM: These neural network based load forecasters are vulnerable to stealthy adversarial attacks!

Motivation for the TeSER Testbed

- This testbed targets evaluation of potential vulnerabilities and successful resilient strategies for complex Cyber-Physical Systems (initially in the power-grid domain). Although adversarial machine learning is not new, it's application in the context of security and resilience of CPS is novel.
- In power-grid, traditional load forecasting doesn't work well with highly dynamic variations in smart grid topology (e.g., bidirectional power flow) and power supply and demand (e.g., DERs and time of use rate). Here, it would require updating models continuously, which is not practical.
- Machine-learning methods can handle these, but suffer from black-box problem. Also, modern grid now has much higher digital connectivity among grid and control equipment. These two makes ML methods susceptible to adversarial attacks. So, we need a testbed that can help with evaluating vulnerabilities and successful resilient strategies.

- Another key motivation for this testbed is to support a web-based, collaborative, model-based approach that can enable rapid prototyping and experimentation with various neural network architectures and data processing, training and evaluation pipelines.
- TeSER also aims to support tight integration with the CPS simulation tools – such as GridLAB-D – which further simplifies the process and shortens the time for input data generation for such models.
- **Note:** All of the testbed tools and technologies are largely domain-independent. So, these methods can be directly utilized in other CPS application domains such as transportation, biomedical, defense, etc.

TeSER: Testbed for Simulation-Based Evaluation of Resilience



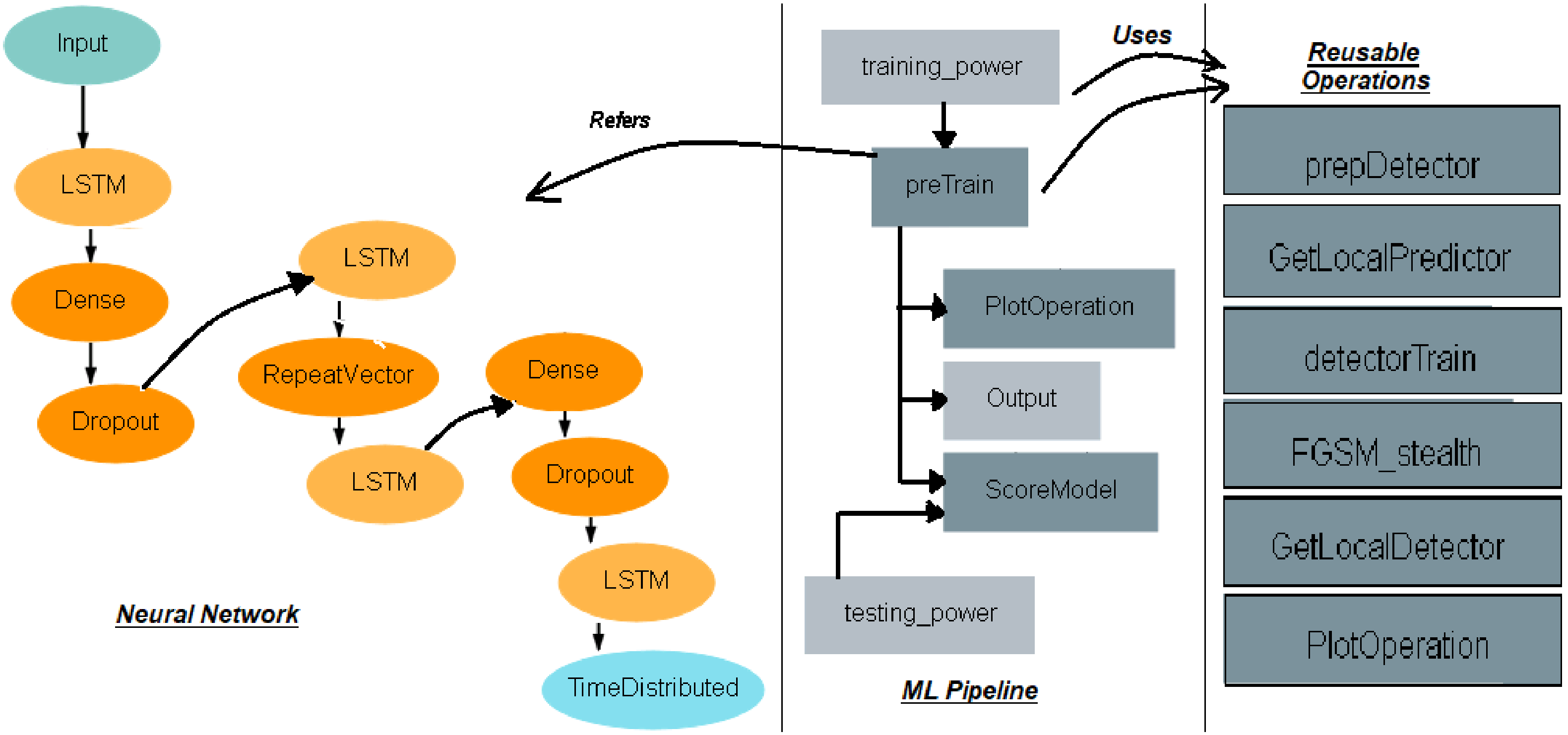
TeSER Testbed Architecture

Web-accessible, Collaborative, Cloud-Supported

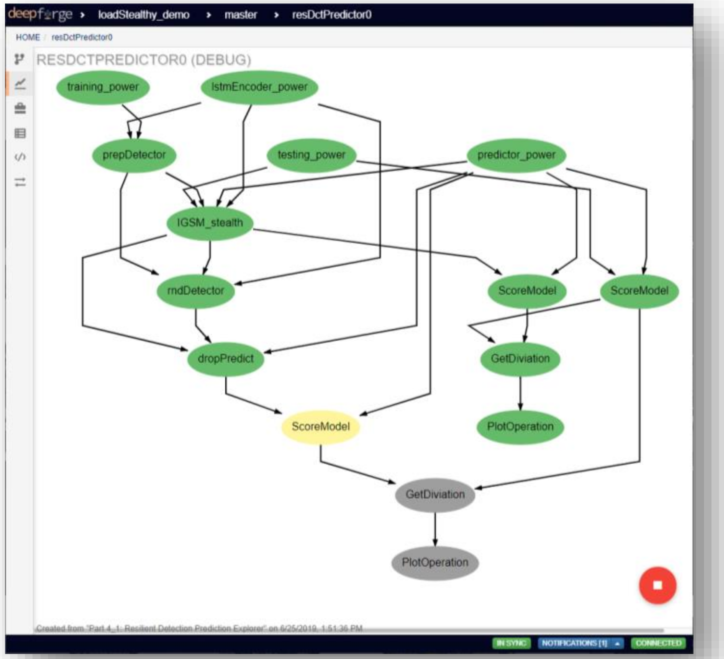
TeSER Testbed (requires password)
https://lablet.webgme.org

- Built using four "open-source" technologies:
  - **WebGME** (Web-based Generic Modeling Environment): Meta-modeling environment for creating rich domain-specific modeling languages
  - **GridLAB-D**: Power grid distribution systems steady-state simulator
  - **DeepForge**: Deep Learning Framework
  - **MongoDB**: Object-oriented database
- Integrated cloud computation platform for executing large-scale experiments
- Integrated support for modeling various Tensorflow/ Keras based machine learning architectures
- Supports storage of experiment results and presenting as digestible plots
- Full versioning and change-tracking of all models
- Full record of executed ML pipelines: iterations, console logs, etc.

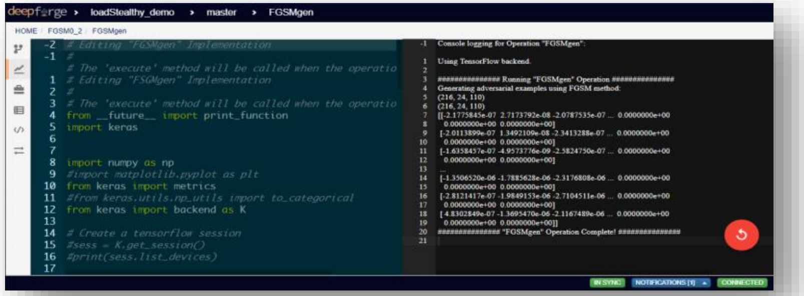
Deep Learning Framework



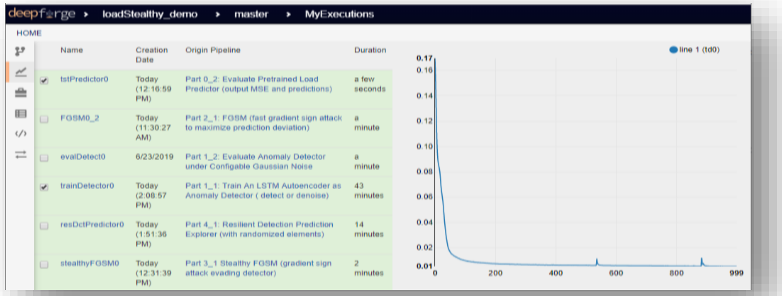
Pipeline Execution



Code editor and console output view



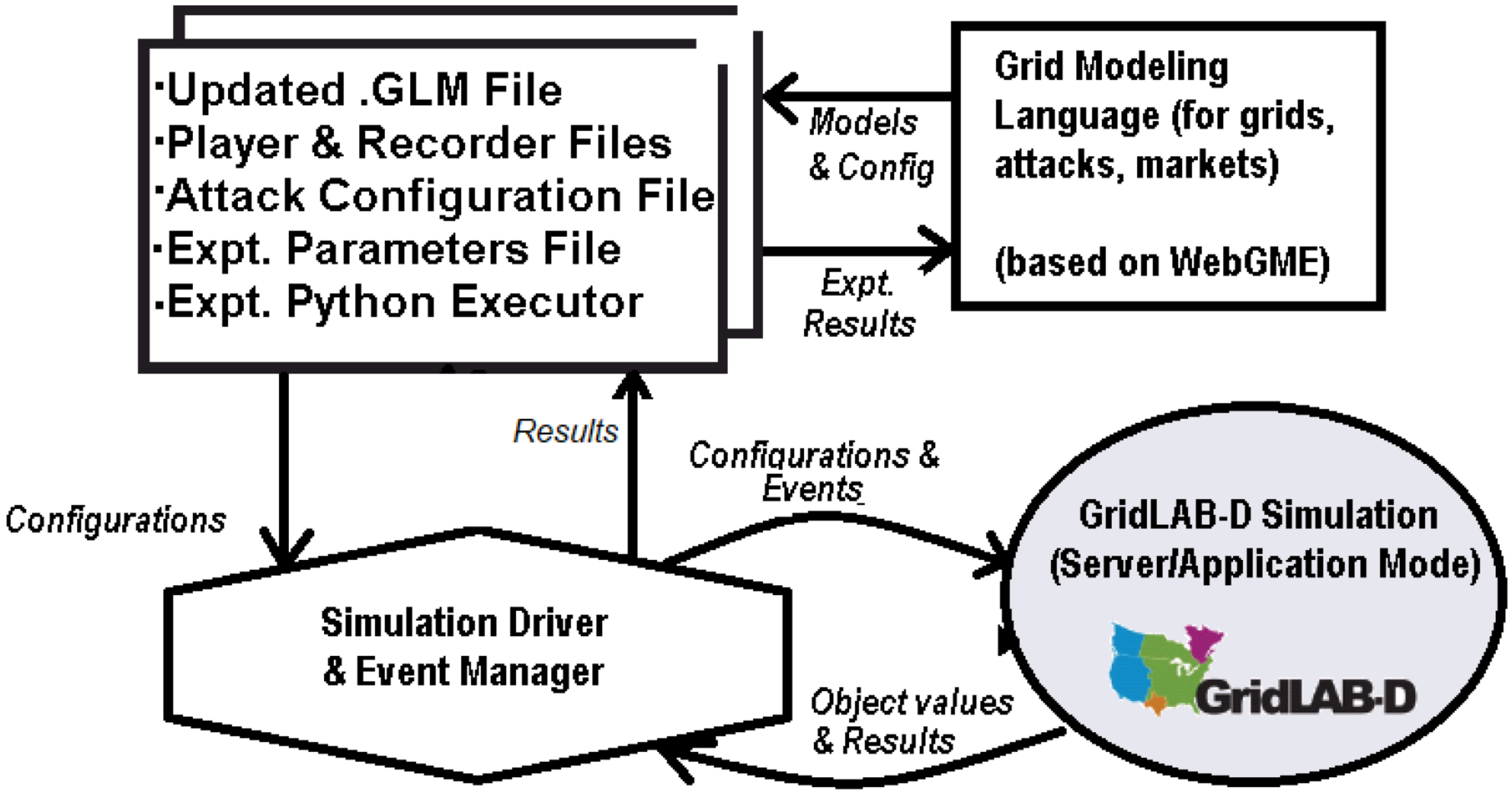
Integrated plotting of executions



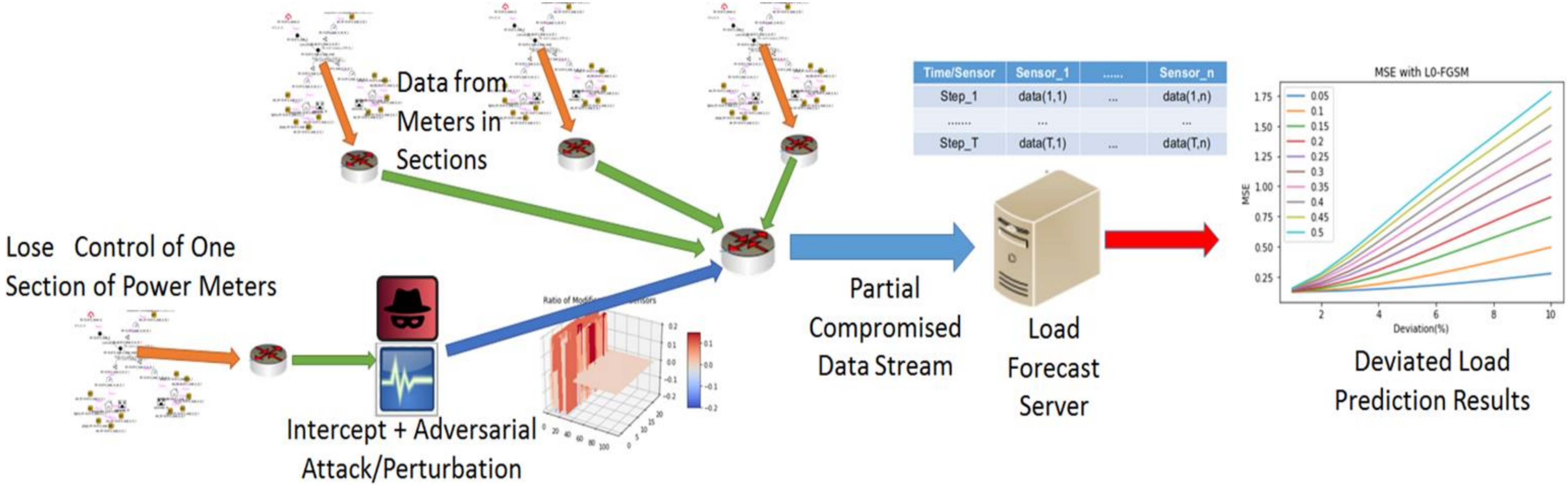
VANDERBILT UNIVERSITY

NIST National Institute of Standards and Technology

Distribution Grid M&S

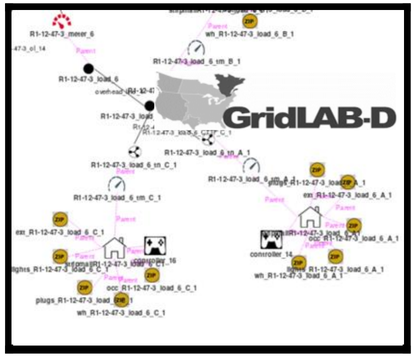


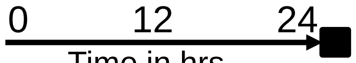
Evaluating Adversarial Attack Impact on Grid Forecasters



Ex 1: Comparing Deep Learning Based Load Predictors

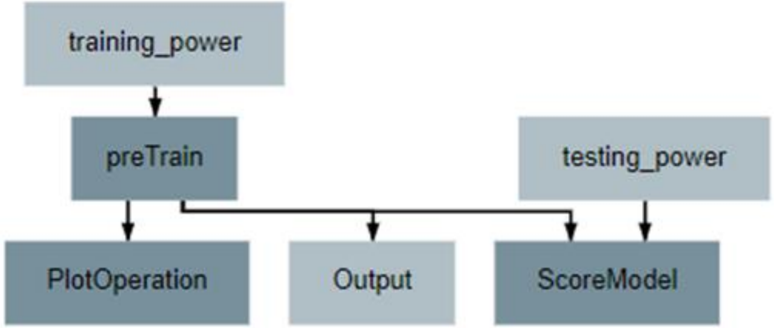
Medium scale feeder in GridLAB-D (109 smart meters)

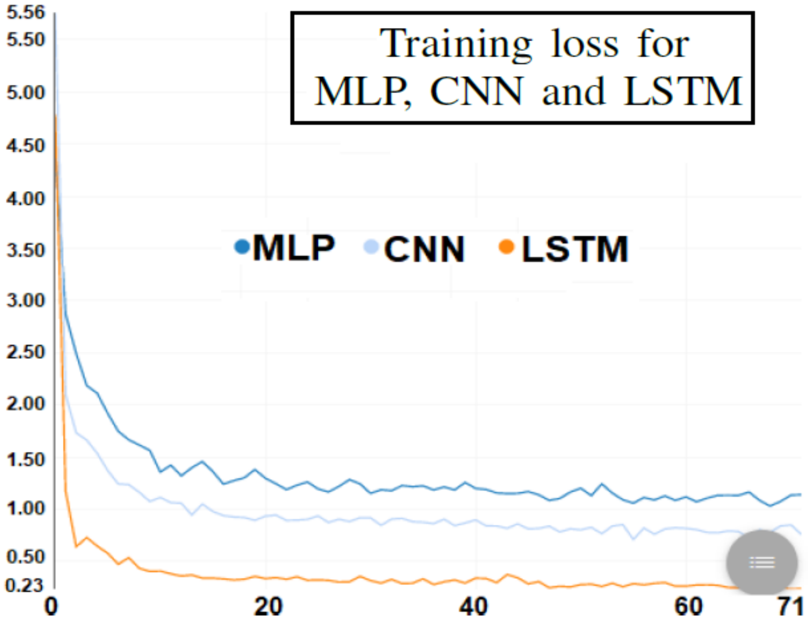




Prediction hour

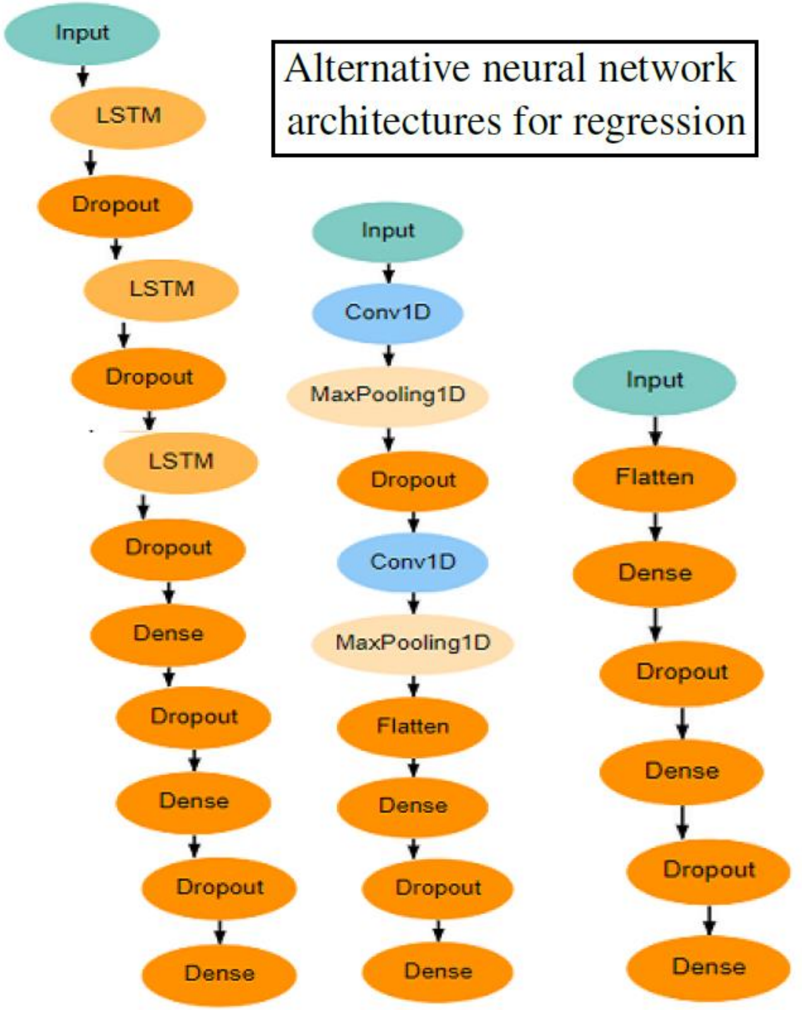
Pipeline model for training predictors





Training loss for MLP, CNN and LSTM

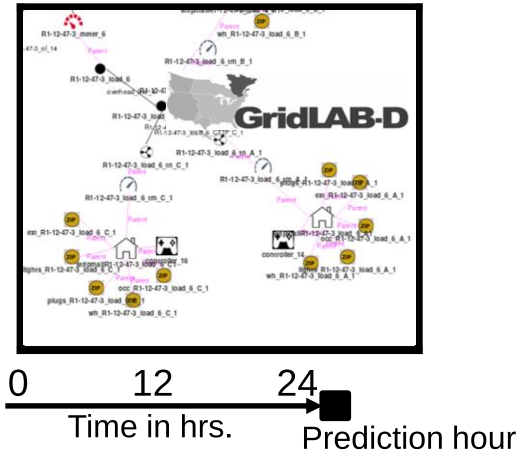
Alternative neural network architectures for regression



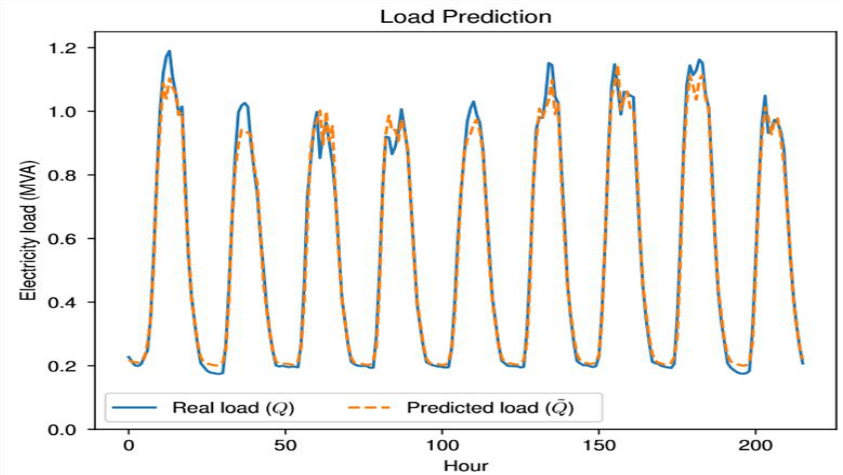
LSTM model CNN model MLP model

Ex2: Load Predictions under Stealthy Adversarial Attacks

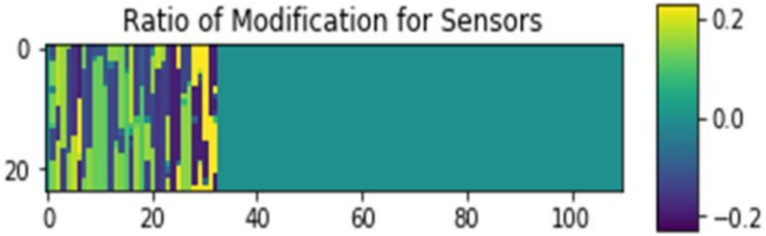
- Medium scale feeder in GridLAB-D (109 smart meters)



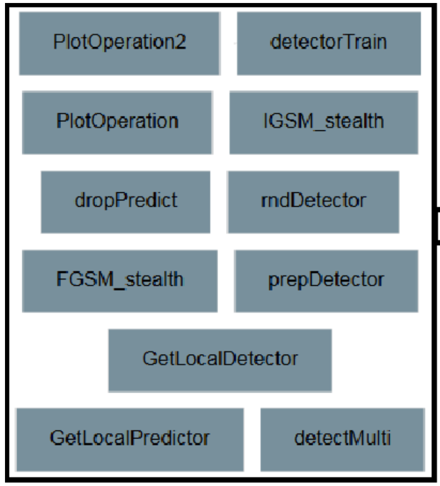
- LSTM load forecast predictor
- Auto-encoder anomaly detector



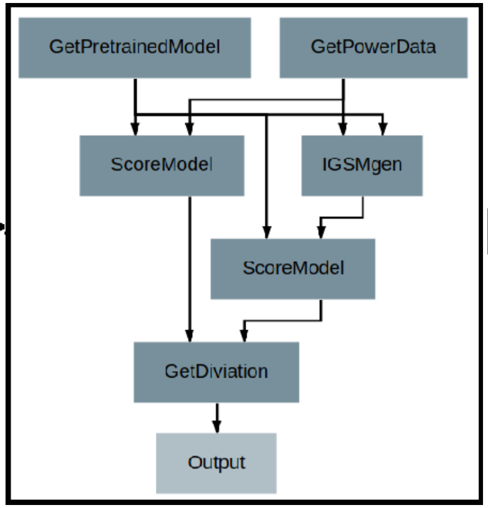
- Threat constraints: 30% of sensors compromised, each modified no more than 20%
- Assume worst-case **white-box** attacks (i.e., full knowledge of predictor and anomaly detector)



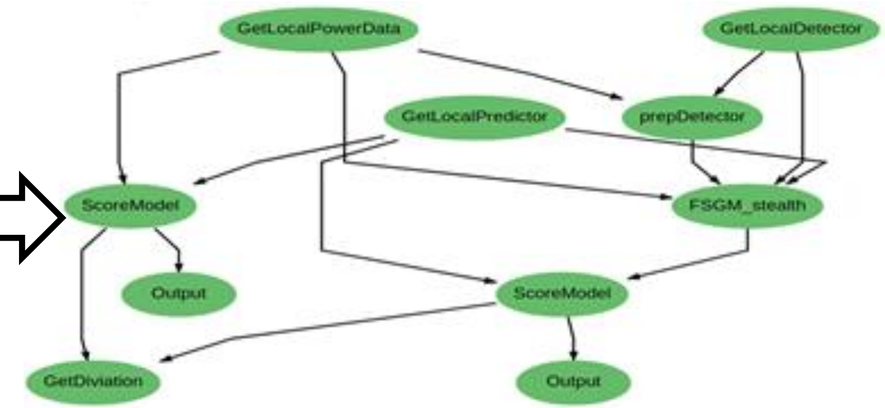
Reusable **Operations** in TeSER



ML **Pipelines** in TeSER



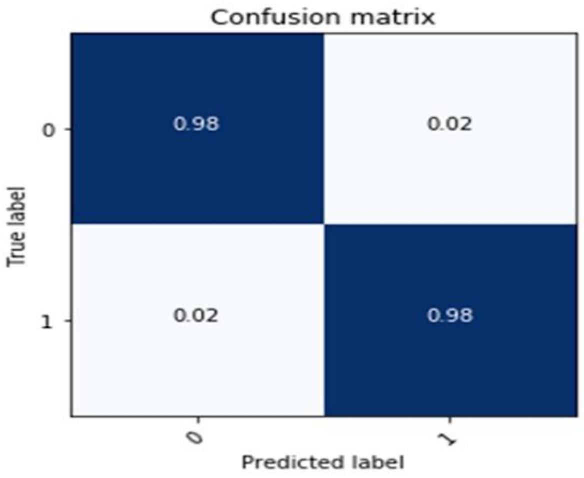
ML Pipelines “**Executions**” in TeSER

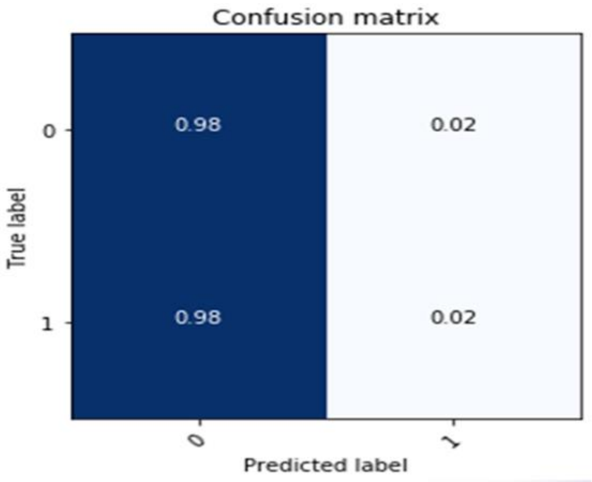


Ex2: Experiment Results

Four adversarial attack settings:

- **Fast Gradient Sign Method (FGSM)**: Single step attack to maximize the prediction deviation from the original predictor
- **Iterative GSM**: Iterative attack to maximize the prediction deviation from the original predictor
- **Directed GSM** (reverse = 1): Iterative attack to minimize the predicted values
- **Directed GSM** (reverse = −1): Iterative attack to maximize the predicted values





Prediction Results (MSE) with Different Prediction Deployment Settings

Attack/Detection Settings	Original/NoAttack	Adversarial/NoDetect	Original/StaticDetect	Adversarial/StaticDetect
Fast-GSM (rate=0.3,step_len=0.2)	0.1255	0.5375	0.1287	0.5322
Iterative-GSM (rate=0.3, step_len=0.01,step_num=20)	0.1255	0.7801	0.1287	0.7606
DirectedGSM (rate=0.3, step_len=0.01 ,step_num=20, reverse=1)	0.1255	0.4785	0.1287	0.4913
DirectedGSM (rate=0.3, step_len=0.01 ,step_num=20, reverse=-1)	0.1255	1.025	0.1287	0.9899

References

[1] Borges, Cruz E., et al., "Evaluating combined load forecasting in large power systems and smart grids," *IEEE Trans. on Industrial Informatics*, 2012, *doi*:10.1109/TII.2012.2219063.

[2] McDaniel, Patrick, et al., "Machine learning in adversarial settings," *IEEE Security & Privacy*, 2016, *doi*:10.1109/MSP.2016.51.

[3] Kurakin, Alexey, et al., "Adversarial machine learning at scale," 2016, *arXiv*:1611.01236v2.

[4] Ghafouri, Amin, et al., "Adversarial regression for detecting attacks in cyber-physical systems," 2018, *arXiv*:1804.11022v1.

[5] S. Soltan, M. Yannakakis, and G. Zussman, "React to cyber-physical attacks on power grids," *ACM SIGMETRICS Performance Evaluation Review*, vol. 46, no. 2, pp. 50–51, 2019, *doi*:10.1109/TNSE.2018.2837894.

[6] "Testbed for Simulation-based Evaluation of Resilience (TeSER)," February 2020, *URL*:http://lablet.webgme.org.

[7] Kecskés, Tamás, et al., "Bridging engineering and formal modeling: Webgme and formula integration." in *MODELS (Satellite Events)*, 2017, *Semanticscholar*:44617122.

[8] Broll, Brian, et al., "Deepforge: A scientific gateway for deep learning," *Gateways*, 2018, *doi*:10.6084/m9.figshare.7092272.v2.

[9] Chassin, David P., et al, "GridLAB-D: An open-source power systems modeling and simulation environment," in *IEEE PES T&D Conference and Exposition, 2008*, pp. 1–5, *doi*:10.1109/TDC.2008.4517260.

[10] Zhou, Xingyu, et al., "Evaluating Resilience of Grid Load Predictions under Stealthy Adversarial Attacks," *Resilience Week Symposium*, 2019, *doi*:10.1109/RWS47064.2019.8971816.

[11] N. Phuangpornpitak and W. Prommee, "A study of load demand forecasting models in electric power system operation and planning," *GMSARN Int. Journal*, 2016, *Semanticscholar*:52992311.

[12] H. He and J. Yan, "Cyber-physical attacks and defences in the smart grid: a survey," *IET Cyber-Physical Systems: Theory & Applications*, vol. 1, no. 1, pp. 13–27, 2016, *doi*:10.1049/iet-cps.2016.0019.

[13] Deka, Deepjyoti, et al., "Optimal data attacks on power grids: Leveraging detection & measurement jamming," in *2015 IEEE SmartGridComm*, pp. 392–397, *doi*:10.1109/SmartGridComm.2015.7436332.

[14] Neema, Himanshu, et al., "Web-Based Platform for Evaluation of Resilient and Transactive Smart-Grids," in *MSCPES 2019*. IEEE, 2019, pp. 1–6, *doi*:10.1109/MSCPES.2019.8738796.