

# 2021-10-22 AI/ML for NFV Meeting Minutes

## Attendees

Beth Cohen

Girish L

Sridhar Rao

Ildiko Vancsa

Al Morton

Sl. No.	Topic	Presenter	Notes
1	Model-Selector Update	Sridhar Rao	<a href="https://github.com/opnfv/thoth/tree/master/tools/modelselector">https://github.com/opnfv/thoth/tree/master/tools/modelselector</a>  Patch is submitted.  Source of the flowchart is also included. .drawio file. Can be modified with <a href="https://app.diagrams.net/">https://app.diagrams.net/</a> , or its desktop version  <b>ToDo:</b> Create documentation, more testing.
2	FP Model Dev. Update		Patch is merged.  <a href="https://github.com/opnfv/thoth/tree/master/models/failure_prediction/jnotebooks">https://github.com/opnfv/thoth/tree/master/models/failure_prediction/jnotebooks</a>  Orange Data: <a href="https://www.kaggle.com/imenbenyahia/clearwatervnf-virtual-ip-multimedia-ip-system">https://www.kaggle.com/imenbenyahia/clearwatervnf-virtual-ip-multimedia-ip-system</a>  <b>ToDo:</b> Documentation.  Note: There will be no progress on this work going ahead until we get a fresh data (maybe a different failure). Desperately need a common data model. <a href="#">Beth Cohen</a> will talk to the EUAG about actually working on this project of creating the data model. Also talk to <a href="#">Steven Casey</a> about anonymizing some VZ data to share. Not sure what the status is of the state of the data.  Question what type of data are we looking for? <ul style="list-style-type: none"><li>• Look at Orange's anonymized port stats, CPU memory, storage usage. No traffic data is available currently.</li><li>• Can we look at creating traffic flow data? Need to standardize definitions</li></ul>
3	Kubernetes Failure Emulation		<a href="#">Sridhar Rao</a> : going ahead this will be the focus
4	Whitepaper	<a href="#">Beth Cohen</a>	<a href="#">Beth Cohen</a> : Put-together CMCC, Telecom Italia and Verizon.  Key Takeaways: maturity is low, more research and development needed, most of the work in operational and maintenance domain, low-hanging fruit - service assurance issues, shared understanding of intelligent n/w needed, reliable data access is difficult, community can play important role in creating a standardized shared data set. Paper will be published Nov 2021, with a webinar scheduled for Nov. <a href="#">Sridhar Rao</a> is invited to participate in Webinar.
5	Tool: Data Extraction		No Progress.
6	Explore: GANs for Synthetic Data Generation		Work in progress. Next week (29th Oct), demo of initial metrics-data generation.  <a href="#">Sridhar Rao</a> and <a href="#">Girish L</a> to attend Barometer project and share the work for feedback and getting more reference data.
7	Explore: Openstack Log Analysis with Google BERT		Nova-Logs (Nova-compute, nova-scheduler, and nova-api).  BERT for anomaly detection - work in progress.