Modeling High Performance Computing System Log Messages for Early Prediction of Job Outcome

ALEXANDRA DELUCIA | POST BACCALAUREATE

ELISABETH BASEMAN | MENTOR

USRC/HPC-DES

Los Alamos NATIONAL LABORATORY EST. 1943

Ultrascale Systems Research Center

LA-UR-18-27491

Motivation

- Predict job failure
  - Help users and system admins
- Research semi/unsupervised HPC log analysis tools
  - Approaching exascale computing
  - Syslog analysis techniques can be transferrable to other tools

Research Questions

1. How accurately can the outcome of a job be predicted using system logs?

2. Which features from system logs work the best?

3. How early can we predict job outcome?

Outline

- Background: Job Logs, Syslogs, and Machine Learning, Oh My!
- Syslog Feature Extraction
- Phase 1: Predicting Job Outcome
- Phase 2: Early Prediction of Job Outcome
- Summary
- Applications and Future Work

Job Logs

- Job: allocation of resources assigned to a user for a specified amount of time[4]
  - i.e. memory, processing power
  - Runs on a cluster such as Grizzly, Wolf, Darwin
- Jobs are recorded by the job scheduler in a **job log file**
  - e.g. Moab, Slurm

```
JobID=# UserID=# GroupID=# Name=<program name>
JobState=[COMPLETED,FAILED,NODE_FAIL,CANCELLED,TIMEOUT] Partition=<>
TimeLimit=# StartTime=<time> EndTime=<time> NodeList=[] NodeCnt=#
ProcCnt=# WorkDir=../../
```

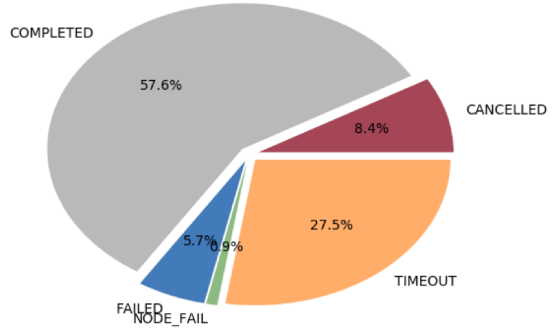
Job log entry format

Job State Frequency

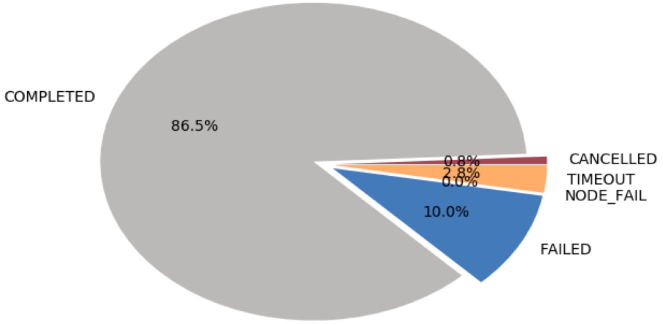
Job State	Description	Okay vs. Problem
Cancelled*	User cancelled the job	Okay
Completed	Job completed successfully	Okay
Failed	Job did not complete for some reason (e.g. program bug)	Problem
Node Fail	One or more of the job's compute nodes failed (e.g. filesystem error)	Problem
Timeout	Job did not finished in the allocated time limit	Okay

*The “cancelled" job state is not used in our experiment

Wolf



Grizzly



System Logs (Syslogs)

- Syslog: log file of recorded events from a computer
  - Every node outputs log file lines and they are combined into a single log file
- Gives insight to process completions/failures and aids in computer diagnostics

```
<Datetime> <Node> <Process Tag> <Message>

Mar 26 03:45:02 wf001 TEMP_SENSORS: coretemp +27.0°C
```

Example syslog line

Data Origin

The data was collected from Grizzly and Wolf over different time periods

	Grizzly	Wolf
Number of Compute Nodes	1,490	616
Scheduler	Slurm	Moab
Time Frame	July 5-18 2018	Mar 26-30 2017
Number of Jobs	6,637	1,775
Number of Matching Syslog	1,939,503	1,074,157

Machine Learning | Supervised

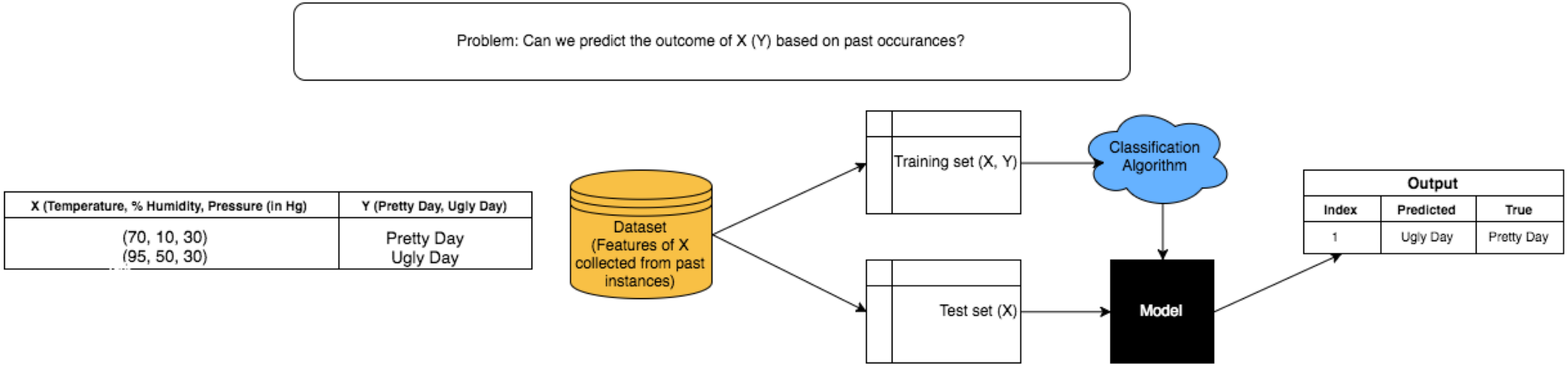
THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS

Using past data and known outcomes to predict outcomes from new data

- e.g. Linear Regression, Random Forests, Neural Networks
- We use Random Forest, a group of decision trees

Machine Learning| Process

1. Gather a set of input **features** from the dataset
   ◦ Input is provided with **labels** (i.e. classes)
2. Partition the dataset into a **train** set and **test** set
3. Run the training set through classification algorithm to create a **model**
4. Evaluate the model's prediction performance on the test set

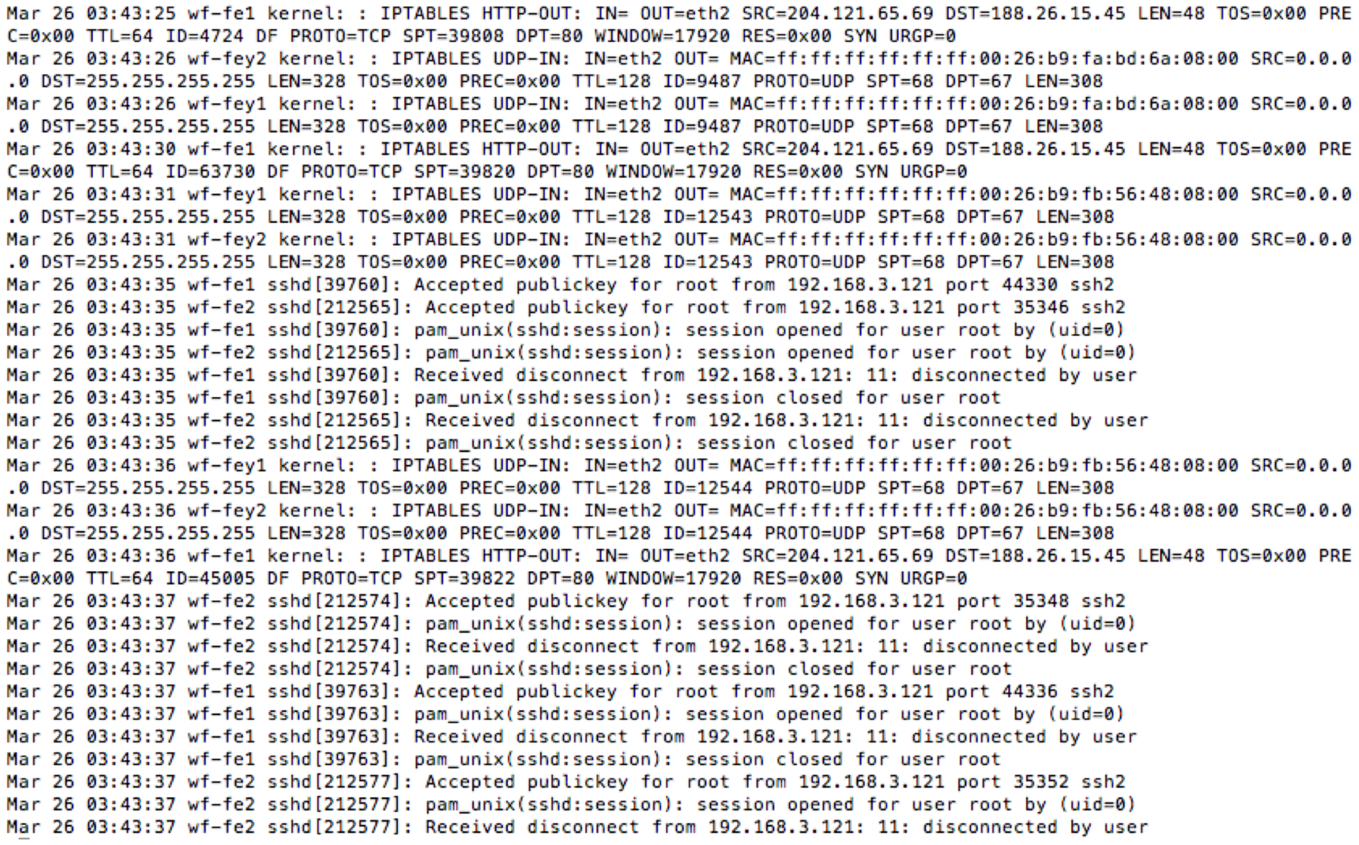


Problem: Can we predict the outcome of X (Y) based on past occurances?

X (Temperature, % Humidity, Pressure (in Hg)	Y (Pretty Day, Ugly Day)
(70, 10, 30) (95, 50, 30)	Pretty Day Ugly Day

Output		
Index	Predicted	True
1	Ugly Day	Pretty Day

Problem: Model Cannot Accept Raw Syslogs as Input



```
Mar 26 03:43:25 wf-fe1 kernel: : IPTABLES HTTP-OUT: IN= OUT=eth2 SRC=204.121.65.69 DST=188.26.15.45 LEN=48 TOS=0x00 PRE
C=0x00 TTL=64 ID=4724 DF PROTO=TCP SPT=39808 DPT=80 WINDOW=17920 RES=0x00 SYN URGP=0
Mar 26 03:43:26 wf-fey2 kernel: : IPTABLES UDP-IN: IN=eth2 OUT= MAC=ff:ff:ff:ff:ff:ff:00:26:b9:fa:bd:6a:08:00 SRC=0.0.0
.0 DST=255.255.255.255 LEN=328 TOS=0x00 PREC=0x00 TTL=128 ID=9487 PROTO=UDP SPT=68 DPT=67 LEN=308
Mar 26 03:43:26 wf-fey1 kernel: : IPTABLES UDP-IN: IN=eth2 OUT= MAC=ff:ff:ff:ff:ff:ff:00:26:b9:fa:bd:6a:08:00 SRC=0.0.0
.0 DST=255.255.255.255 LEN=328 TOS=0x00 PREC=0x00 TTL=128 ID=9487 PROTO=UDP SPT=68 DPT=67 LEN=308
Mar 26 03:43:30 wf-fe1 kernel: : IPTABLES HTTP-OUT: IN= OUT=eth2 SRC=204.121.65.69 DST=188.26.15.45 LEN=48 TOS=0x00 PRE
C=0x00 TTL=64 ID=63730 DF PROTO=TCP SPT=39820 DPT=80 WINDOW=17920 RES=0x00 SYN URGP=0
Mar 26 03:43:31 wf-fey1 kernel: : IPTABLES UDP-IN: IN=eth2 OUT= MAC=ff:ff:ff:ff:ff:ff:00:26:b9:fb:56:48:08:00 SRC=0.0.0
.0 DST=255.255.255.255 LEN=328 TOS=0x00 PREC=0x00 TTL=128 ID=12543 PROTO=UDP SPT=68 DPT=67 LEN=308
Mar 26 03:43:31 wf-fey2 kernel: : IPTABLES UDP-IN: IN=eth2 OUT= MAC=ff:ff:ff:ff:ff:ff:00:26:b9:fb:56:48:08:00 SRC=0.0.0
.0 DST=255.255.255.255 LEN=328 TOS=0x00 PREC=0x00 TTL=128 ID=12543 PROTO=UDP SPT=68 DPT=67 LEN=308
Mar 26 03:43:35 wf-fe1 sshd[39760]: Accepted publickey for root from 192.168.3.121 port 44330 ssh2
Mar 26 03:43:35 wf-fe2 sshd[212565]: Accepted publickey for root from 192.168.3.121 port 35346 ssh2
Mar 26 03:43:35 wf-fe1 sshd[39760]: pam_unix(sshd:session): session opened for user root by (uid=0)
Mar 26 03:43:35 wf-fe2 sshd[212565]: pam_unix(sshd:session): session opened for user root by (uid=0)
Mar 26 03:43:35 wf-fe1 sshd[39760]: Received disconnect from 192.168.3.121: 11: disconnected by user
Mar 26 03:43:35 wf-fe1 sshd[39760]: pam_unix(sshd:session): session closed for user root
Mar 26 03:43:35 wf-fe2 sshd[212565]: Received disconnect from 192.168.3.121: 11: disconnected by user
Mar 26 03:43:35 wf-fe2 sshd[212565]: pam_unix(sshd:session): session closed for user root
Mar 26 03:43:36 wf-fey1 kernel: : IPTABLES UDP-IN: IN=eth2 OUT= MAC=ff:ff:ff:ff:ff:ff:00:26:b9:fb:56:48:08:00 SRC=0.0.0
.0 DST=255.255.255.255 LEN=328 TOS=0x00 PREC=0x00 TTL=128 ID=12544 PROTO=UDP SPT=68 DPT=67 LEN=308
Mar 26 03:43:36 wf-fey2 kernel: : IPTABLES UDP-IN: IN=eth2 OUT= MAC=ff:ff:ff:ff:ff:ff:00:26:b9:fb:56:48:08:00 SRC=0.0.0
.0 DST=255.255.255.255 LEN=328 TOS=0x00 PREC=0x00 TTL=128 ID=12544 PROTO=UDP SPT=68 DPT=67 LEN=308
Mar 26 03:43:36 wf-fe1 kernel: : IPTABLES HTTP-OUT: IN= OUT=eth2 SRC=204.121.65.69 DST=188.26.15.45 LEN=48 TOS=0x00 PRE
C=0x00 TTL=64 ID=45005 DF PROTO=TCP SPT=39822 DPT=80 WINDOW=17920 RES=0x00 SYN URGP=0
Mar 26 03:43:37 wf-fe2 sshd[212574]: Accepted publickey for root from 192.168.3.121 port 35348 ssh2
Mar 26 03:43:37 wf-fe2 sshd[212574]: pam_unix(sshd:session): session opened for user root by (uid=0)
Mar 26 03:43:37 wf-fe2 sshd[212574]: Received disconnect from 192.168.3.121: 11: disconnected by user
Mar 26 03:43:37 wf-fe2 sshd[212574]: pam_unix(sshd:session): session closed for user root
Mar 26 03:43:37 wf-fe1 sshd[39763]: Accepted publickey for root from 192.168.3.121 port 44336 ssh2
Mar 26 03:43:37 wf-fe1 sshd[39763]: pam_unix(sshd:session): session opened for user root by (uid=0)
Mar 26 03:43:37 wf-fe1 sshd[39763]: Received disconnect from 192.168.3.121: 11: disconnected by user
Mar 26 03:43:37 wf-fe1 sshd[39763]: pam_unix(sshd:session): session closed for user root
Mar 26 03:43:37 wf-fe2 sshd[212577]: Accepted publickey for root from 192.168.3.121 port 35352 ssh2
Mar 26 03:43:37 wf-fe2 sshd[212577]: pam_unix(sshd:session): session opened for user root by (uid=0)
Mar 26 03:43:37 wf-fe2 sshd[212577]: Received disconnect from 192.168.3.121: 11: disconnected by user
```

$$\begin{bmatrix} 0001 & 0.328 & 1.4 & 764 & 5.67 & 0.003 \\ 0002 & 0 & 2 & 40 & 9.05 & 0.1587 \\ 0003 & 0.567 & 2.3 & 234 & 4.23 & 1.0012 \\ 0004 & 0.078 & 0.7 & 293 & 8.08 & 0.9809 \end{bmatrix}$$

Must convert the raw syslog into an input feature vector

Feature Extraction | Numerical and Temporal

NUMERICAL

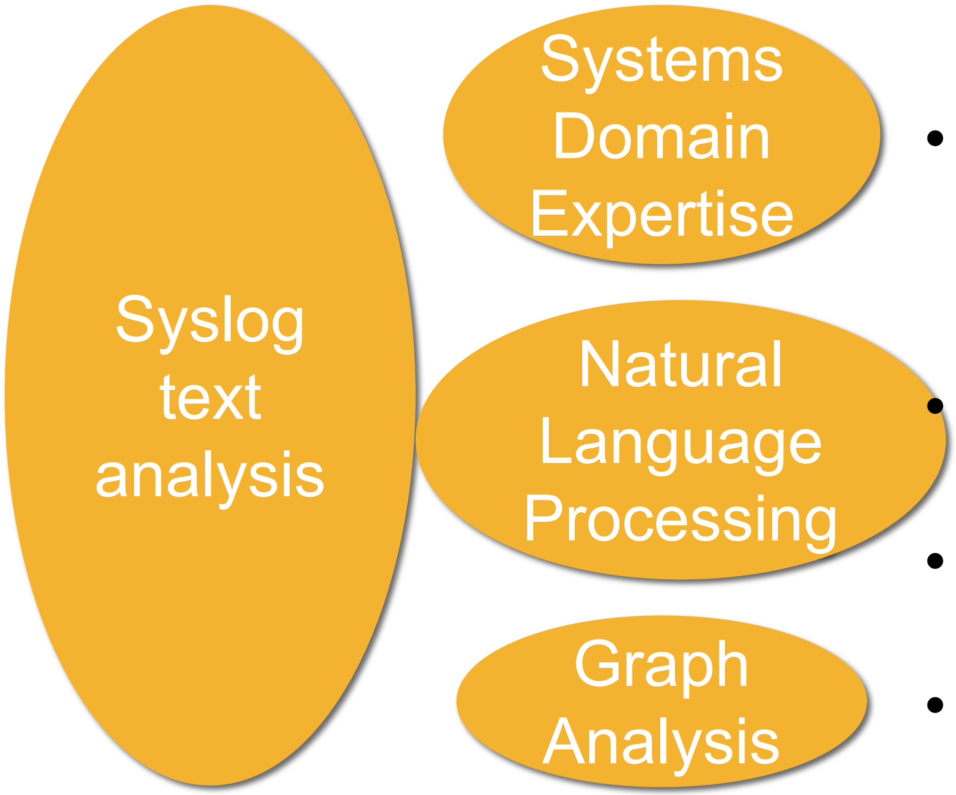
- Average
- Standard deviation
- Count of numbers

TEMPORAL

- Average time between syslog messages
- Standard deviation of time between syslog messages
- Total time between first and last syslog message

Feature Extraction | Text

- Summarize text data using numbers
  - Distributions over clusters or categories
- Techniques originate from different fields



Syslog text analysis

Systems Domain Expertise

Natural Language Processing

Graph Analysis

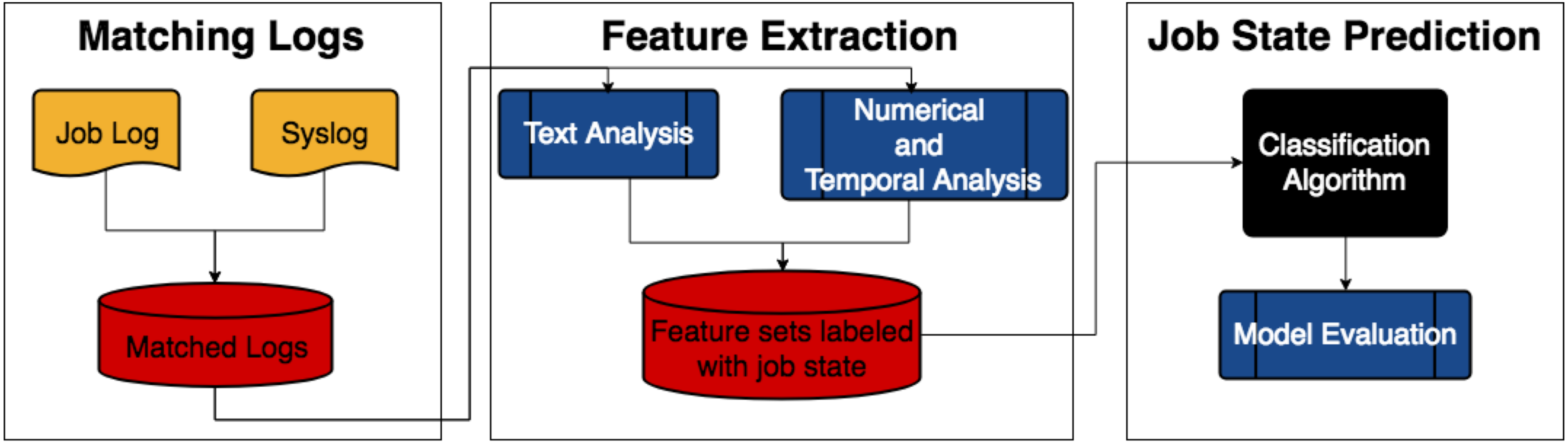
- Tag Clustering
- Topic Model: LDA[1]
- TF-IDF
- Infomap*[2]

*Not used on Grizzly logs

Feature Sets

1) Numerical only
2) Temporal only
3) LDA distribution only
4) Infomap distribution only
5) TF-IDF only
6) Tag distribution only
6) LDA distribution + numerical & temporal
7) Infomap distribution + numerical & temporal
8) TF-IDF + numerical & temporal
9) Tag distribution + numerical & temporal

Methods


Matching Logs
Job Log
Syslog
Matched Logs
Feature Extraction
Text Analysis
Numerical and Temporal Analysis
Feature sets labeled with job state
Job State Prediction
Classification Algorithm
Model Evaluation

Job Outcome Prediction | A Classification Problem

- Job outcomes are described by their corresponding syslogs' text, numerical, and temporal features
- Outcome labels:
  - {COMPLETED, FAILED, NODE_FAIL, TIMEOUT}

Job ID	Job State	Temporal AVG	Temporal STD	Numerical AVG	Numerical STD	Cluster 1	…	Cluster *n*
1	CANCELLED	400	200	30.0	0.05	0.03	……	0.0002

Example of feature table of input for classification algorithm

Job Outcome Prediction | Tasks

Evaluated our feature set models on different tasks

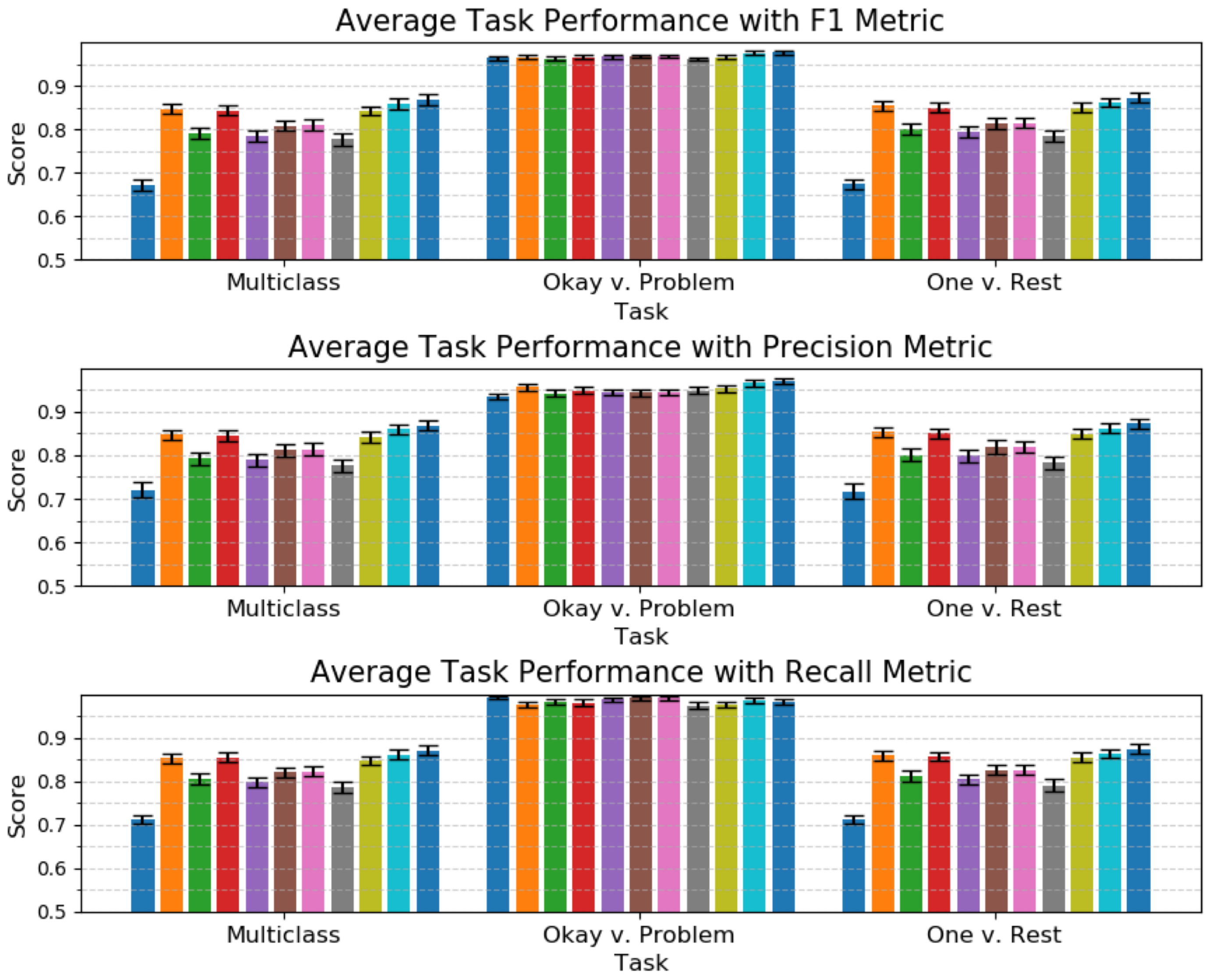
1. Multiclass
   - Predict the job’s outcome from {COMPLETED, FAILED, NODE_FAIL, TIMEOUT}
2. Okay vs. Problem
   - Predict the job’s outcome from {Okay, Problem}
3. One v. Rest
   - One outcome versus the other outcomes {COMPLETED versus FAILED, NODE_FAIL, TIMEOUT}

Job Outcome Prediction | Model Evaluation

- A good model is one that **generalizes** the best
- Precision-Recall metric[5]
  - Measured on [0, 1] where **1 is perfect** and 0.5 is random guessing
- **Precision**: how much of what is returned is correct
  - high precision → low false positive
- **Recall**: how much of what is correct is returned
  - high recall → low false negative
- **F1-score**: harmonic mean of precision and recall

Results | Wolf

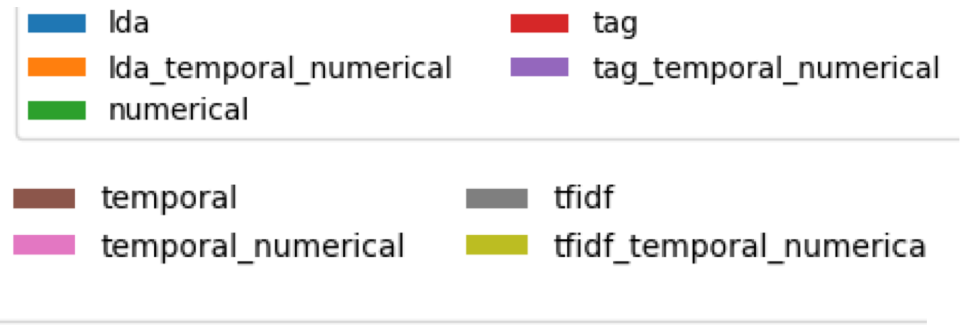
- ■ infomap
- ■ infomap_temporal_numerical
- ■ lda
- ■ lda_temporal_numerical
- ■ numerical
- ■ tag
- ■ tag_temporal_numerical
- ■ temporal
- ■ temporal_numerical
- ■ tfidf
- ■ tfidf_temporal_numerical



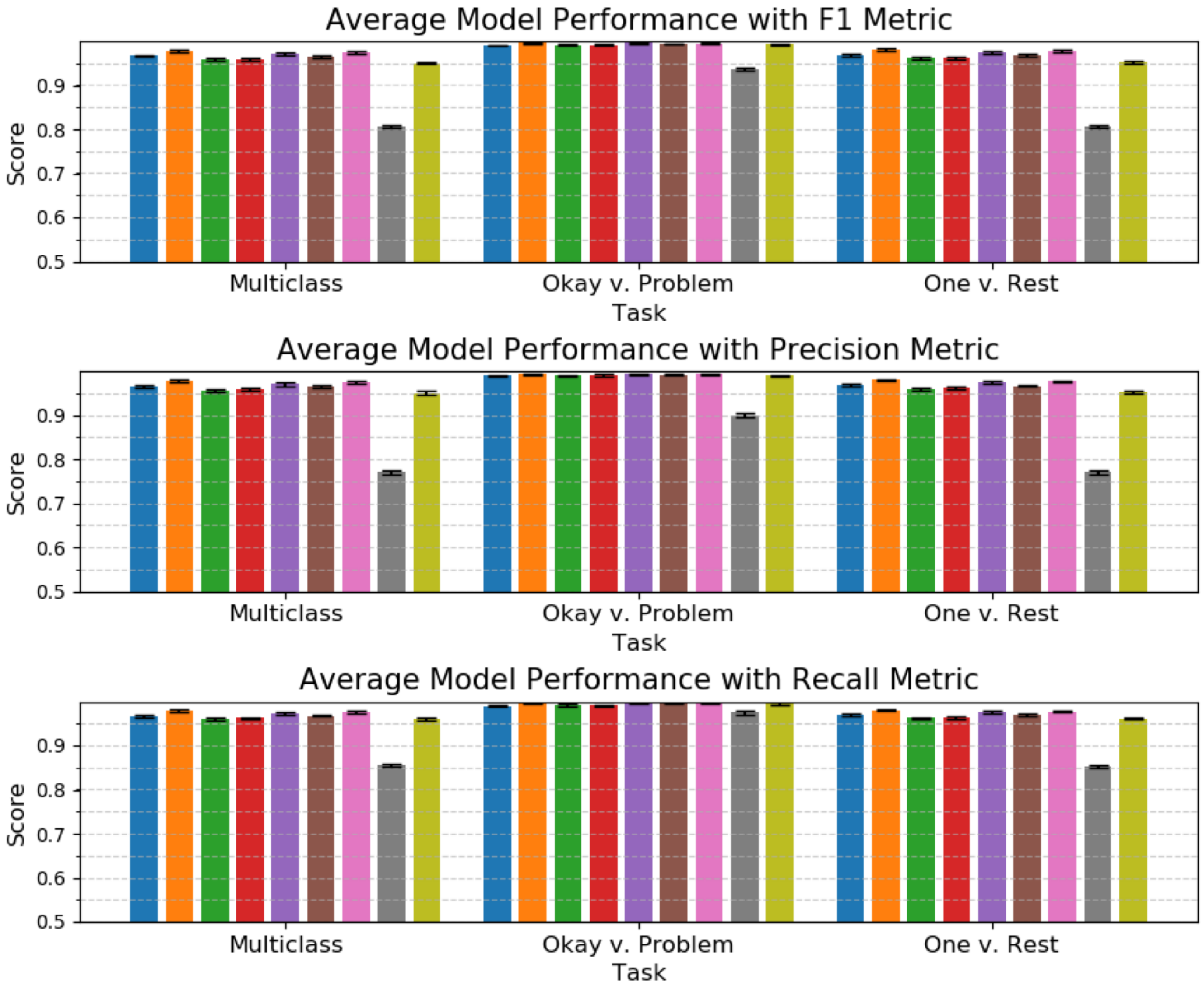
Higher scores are better. Scale starts at 0.5

- The models performed best on the **“Okay v. Problem”** task
- **TF-IDF** feature set performed the best
- **Infomap** feature set performed the worst

Results | Grizzly



- Overall the feature set performances were better than on Wolf
- **LDA** (topic modeling) with temporal and numerical feature set performed the best
- **TF-IDF** feature set performed the worst


Average Model Performance with F1 Metric

Average Model Performance with Precision Metric

Average Model Performance with Recall Metric

Higher scores are better. Scale starts at 0.5

How early can we predict the job outcome?

RETURN TO FINAL RESEARCH QUESTION

Early Prediction

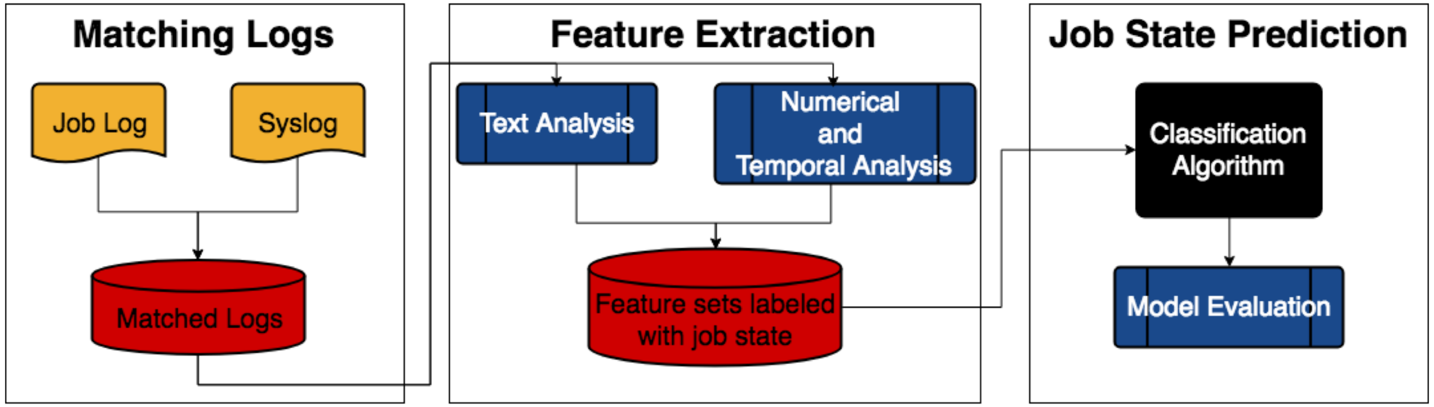
Measure “early” in two ways

1. The **number of syslog messages** into a job
2. **Minutes passed** since the job began

Goal: Real-time log analysis

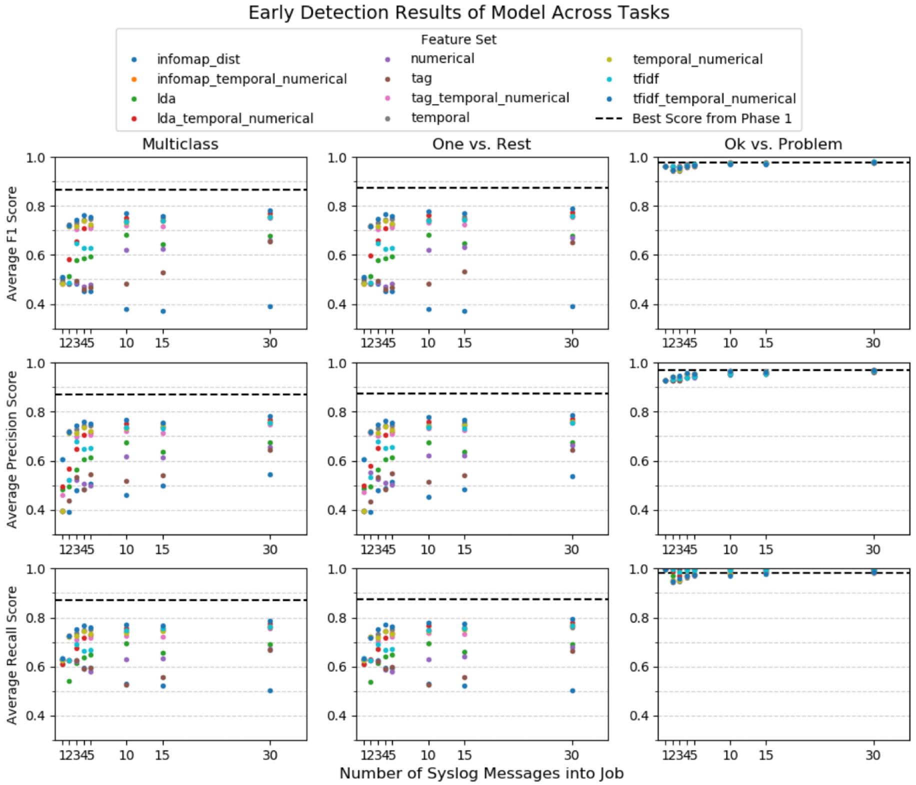
Methods | Early Prediction

1. Match the syslogs to their corresponding jobs that meet the time/message restriction
2. Same process as Phase 1

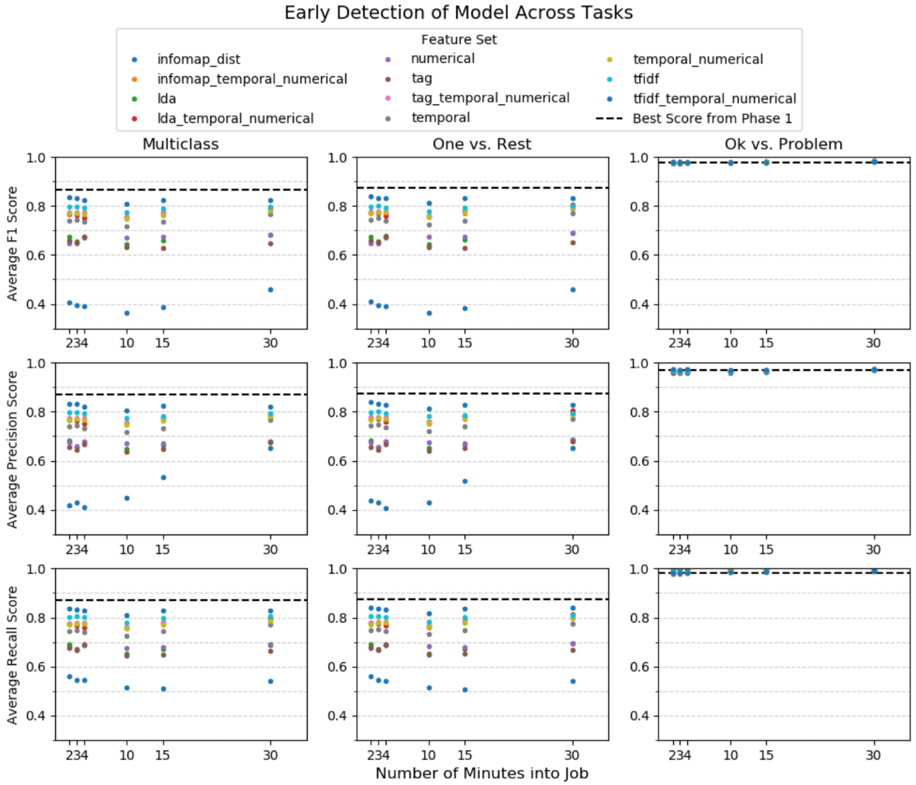


Results | Wolf Early Prediction

- More information → better results
- F1-score of up to 0.7



Limiting Messages



Limiting Minutes

Summary

- Basic features and topics from syslog predict job outcome on two clusters with best F1-scores of above 0.95
- The model trained on the TF-IDF and numerical and temporal features performed the best
- The model was able to predict job outcome with F1 score of over 0.7 when limited to partial Wolf syslog

Applications and Future Work

- Applications
  - Tool to monitor high performance computers and provide real-time predictions for node failure
  - Integrate with a job scheduler for “smart” job checkpointing
- Future Work
  - Train on a larger dataset
  - Test model on different clusters
  - Compare our model to a baseline of current syslog analysis techniques

References

1. Baseman, Elisabeth & Blanchard, Sean & Li, Zongze & Fu, Song. (2016). Relational Synthesis of Text and Numeric Data for Anomaly Detection on Computing System Logs. 882-885. 10.1109/ICMLA.2016.0158.
2. M. Blei, David & Y. Ng, Andrew & Jordan, Michael. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research. 3. 993-1022. 10.1162/jmlr.2003.3.4-5.993.
3. Rosvall, M., Axelsson, D.&Bergstrom, C. Eur. Phys. J. Spec. Top. (2009) 178: 13. https://doi.org/10.1140/epjst/e2010-01179-1
4. SchedMD. Slurm Workload Manager. Web. Accessed 10 Feb 2018.
5. Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.
6. Vaarandi, Risto & Pihelgas, Mauno. (2015). LogCluster - A data clustering and pattern mining algorithm for event logs. 1-7. 10.1109/CNSM.2015.7367331.
7. Wikipedia contributors. Feature extraction. Wikipedia, The Free Encyclopedia. 13 Dec 2017. Web. Accessed 10 Feb 2018.

Questions?