IOWA STATE UNIVERSITY
Department of Civil, Construction
and Environmental Engineering

Crash Identification using Waze Report

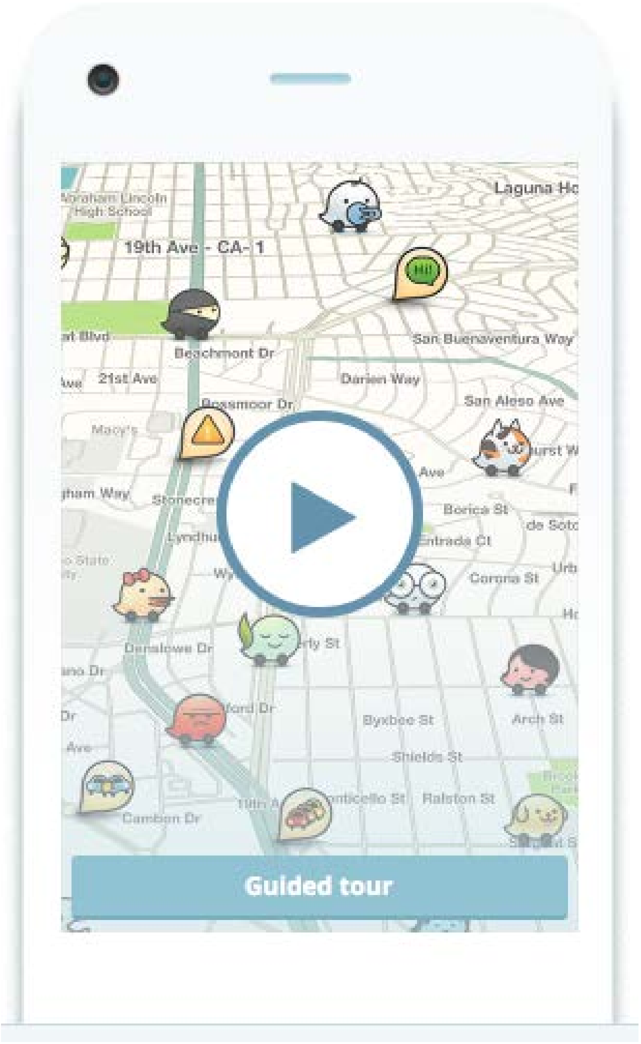
Problem Description – Waze

- Waze is a mobile application for user to report traffic issues

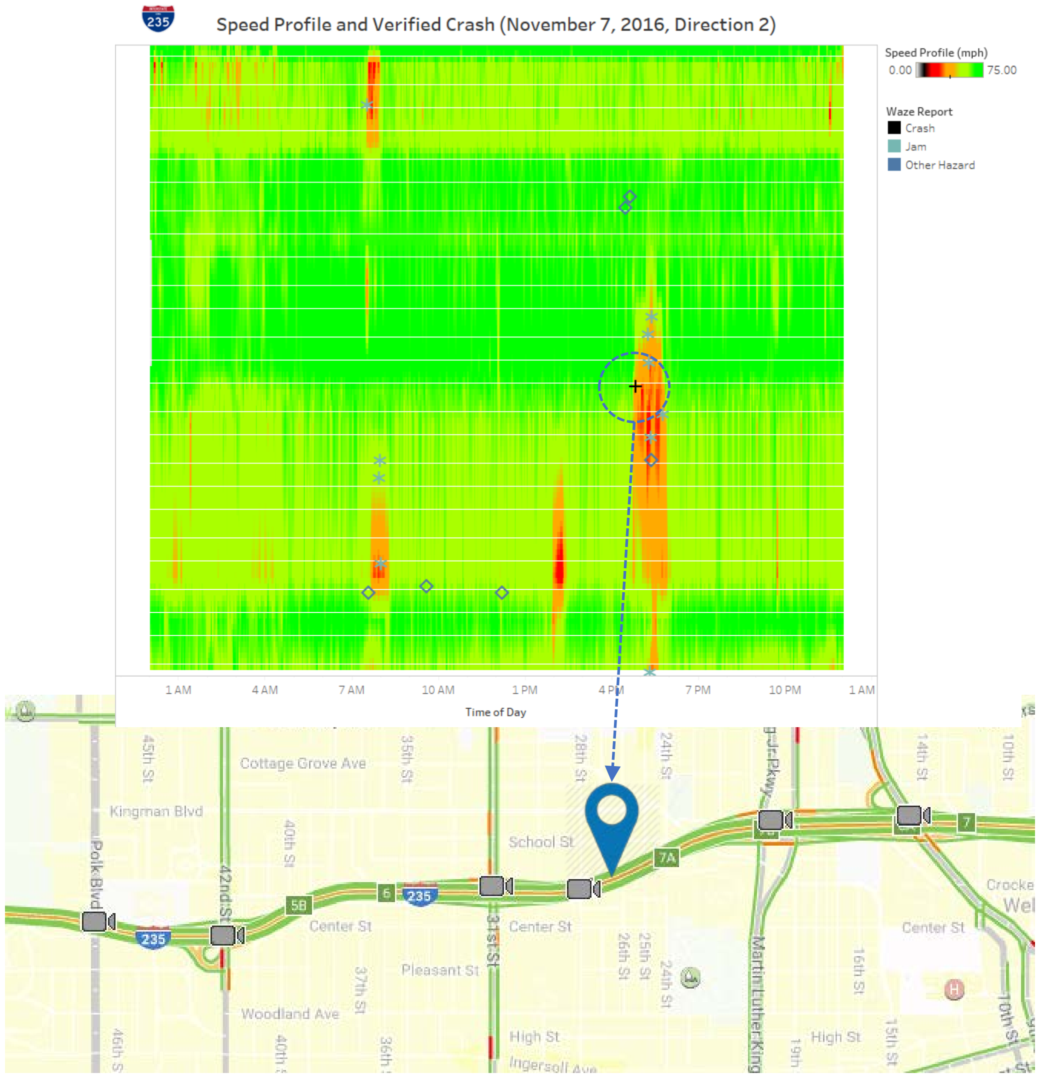
- Three types to report:
  - Accident
  - Jam
  - Hazard

- Google bought Waze in 2013

- Iowa DOT is using Waze for crash identification



Problem Description – Waze



Problem Description - Summary

- Crash or not is what DOT cares about
- It is necessary and beneficial to score Waze reports in terms of crash likelihood (100,000 /year state-wide)
- Waze score their reports based on user credit level
- **<u>We score Waze reports based real traffic and weather condition: P(Crash | traffic, weather)</u>**

Case Study

Location: Interstate 235 in Des Moines, both directions

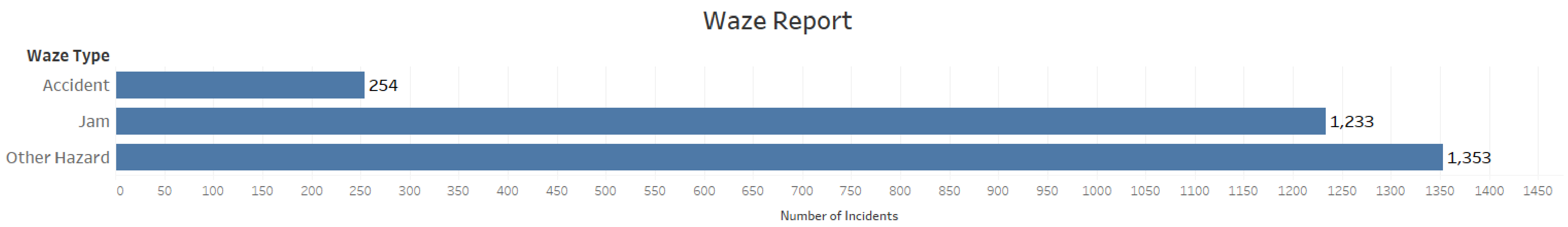
Time Period: Sep. 2016 - Dec.2016

Data Sources:

- Waze reports
- Wavetronix - traffic
- RWIS - weather
- TMC crash reports

Data Description – Waze

- 2,840 – Total
- 254 – Accident
- 1,233 – Jam
- 1,353 – Hazard



Data Description – Wavetronix

- Radar sensors collect traffic data including:

vehicle speed (mi/h);

volume (veh/h);

sensor occupancy (%).

Wavetronix Sensor

- 15 sensors on EB /14 sensors on WB.

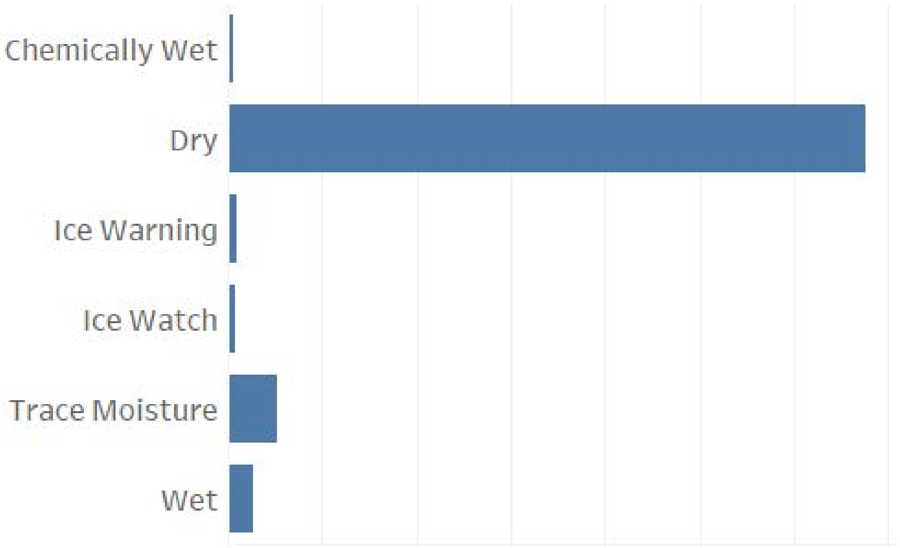
- Data resolution: every 20 seconds.

Sensor Locations

Data Description – RWIS

- Road Weather Information System (RWIS)
  - Temperatures
  - Pavement conditions

- Pavement Conditions:



Dry = 0

Not Dry = 1

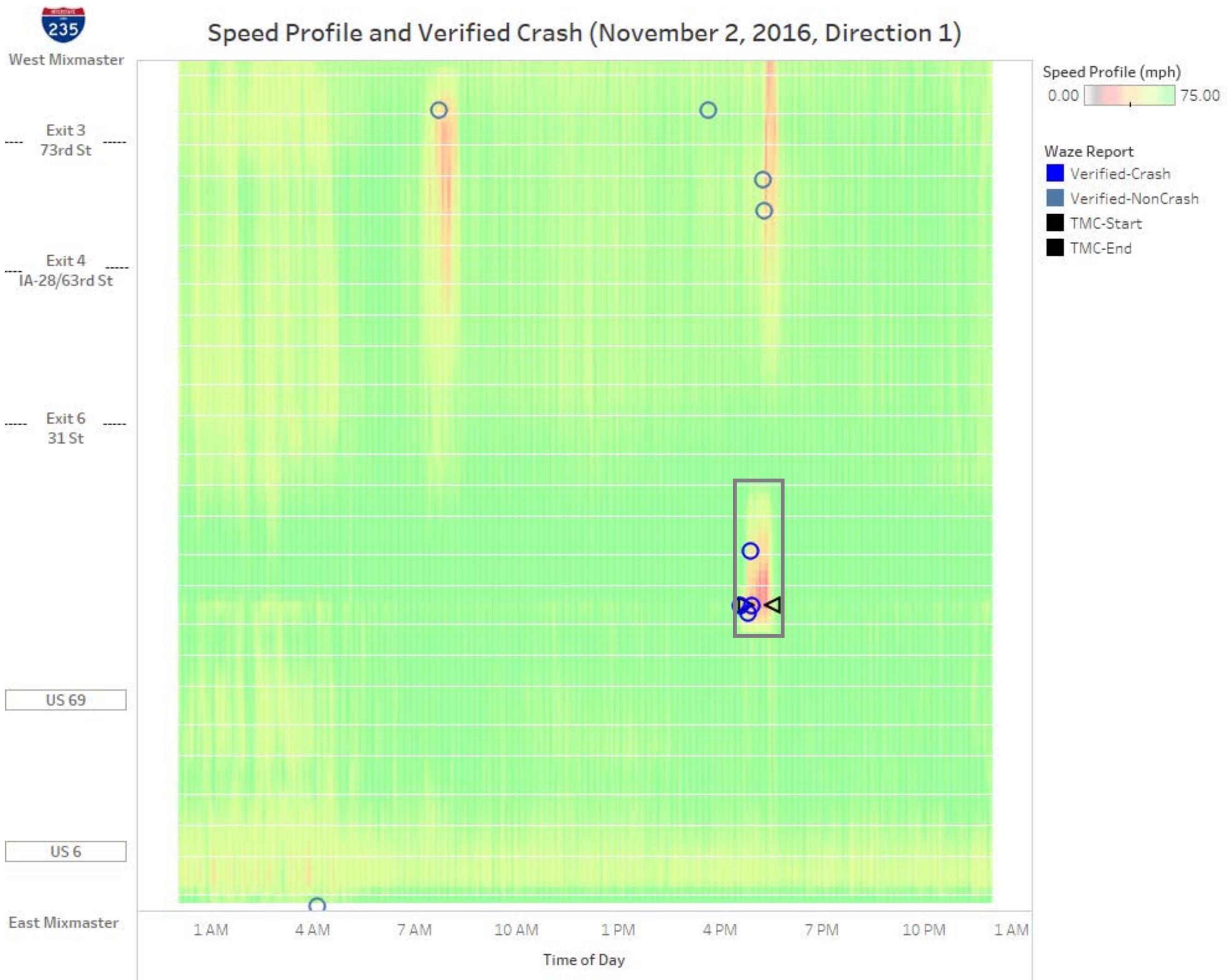
Weather Sensor RWIS Used

Sensor Location in Our Study Area

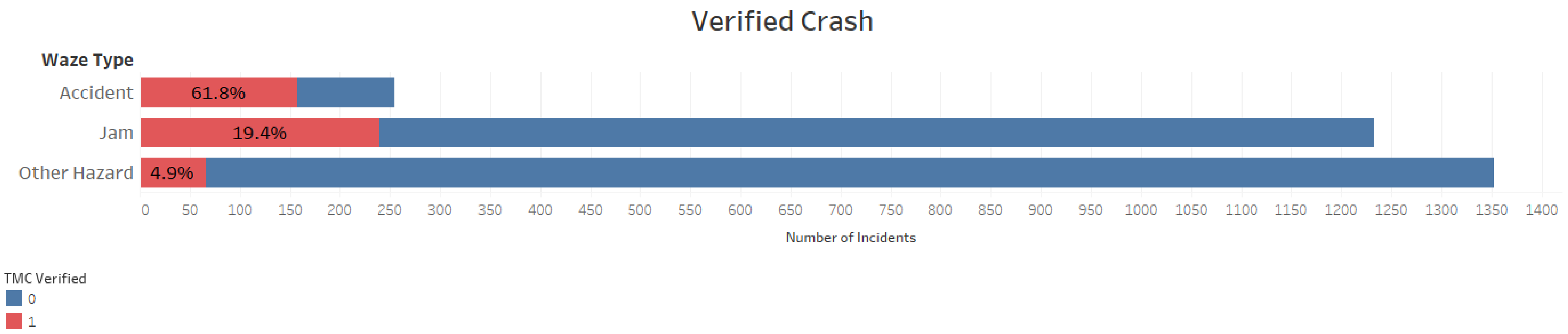
Data Description – TMC crash reports

- TMC = Traffic Management Center of Iowa DOT
- 239 verified crash reports

Data Preparation – Labels



Data Preparation – Label



	ACCIDENT	JAM	HAZARD	Total
Label 1	157	239	66	462
Label 0	97	994	1287	2378
Total	254	1233	1353	2840

Data Preparation – Features

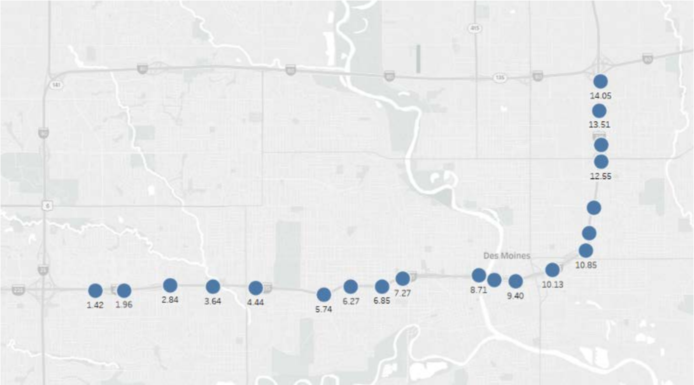
Raw Data → Full Mesh-grid

- Vertical axis

Discrete Sensor Mile Marker → Every 0.1 miles

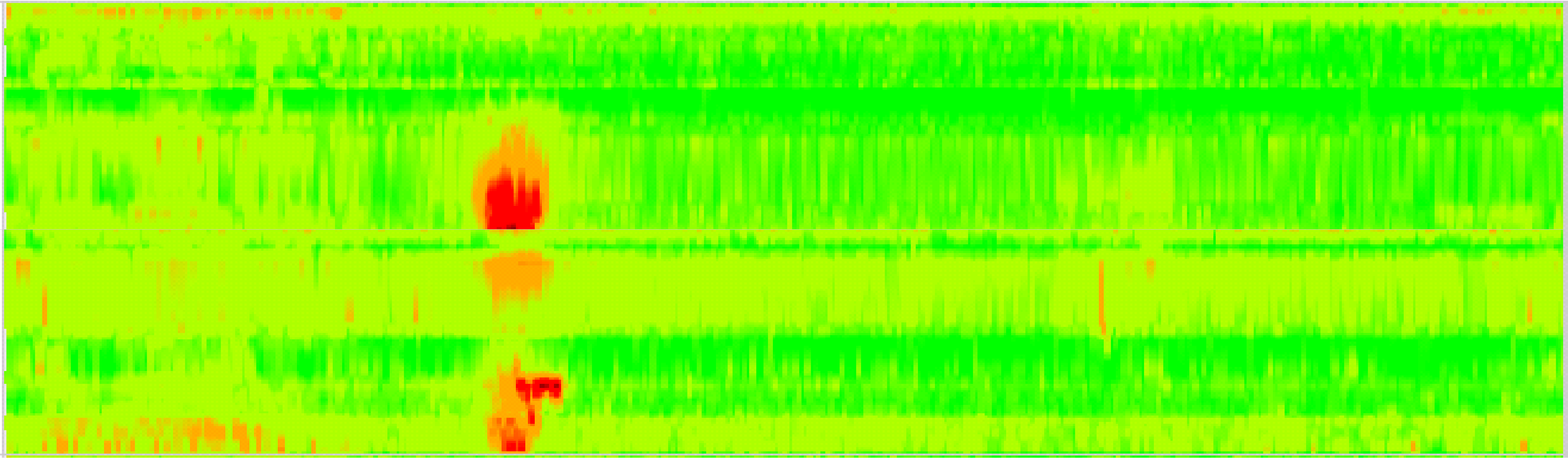
- Horizontal axis

1 Minute with Null Values → Every 1 minute



Eastbound

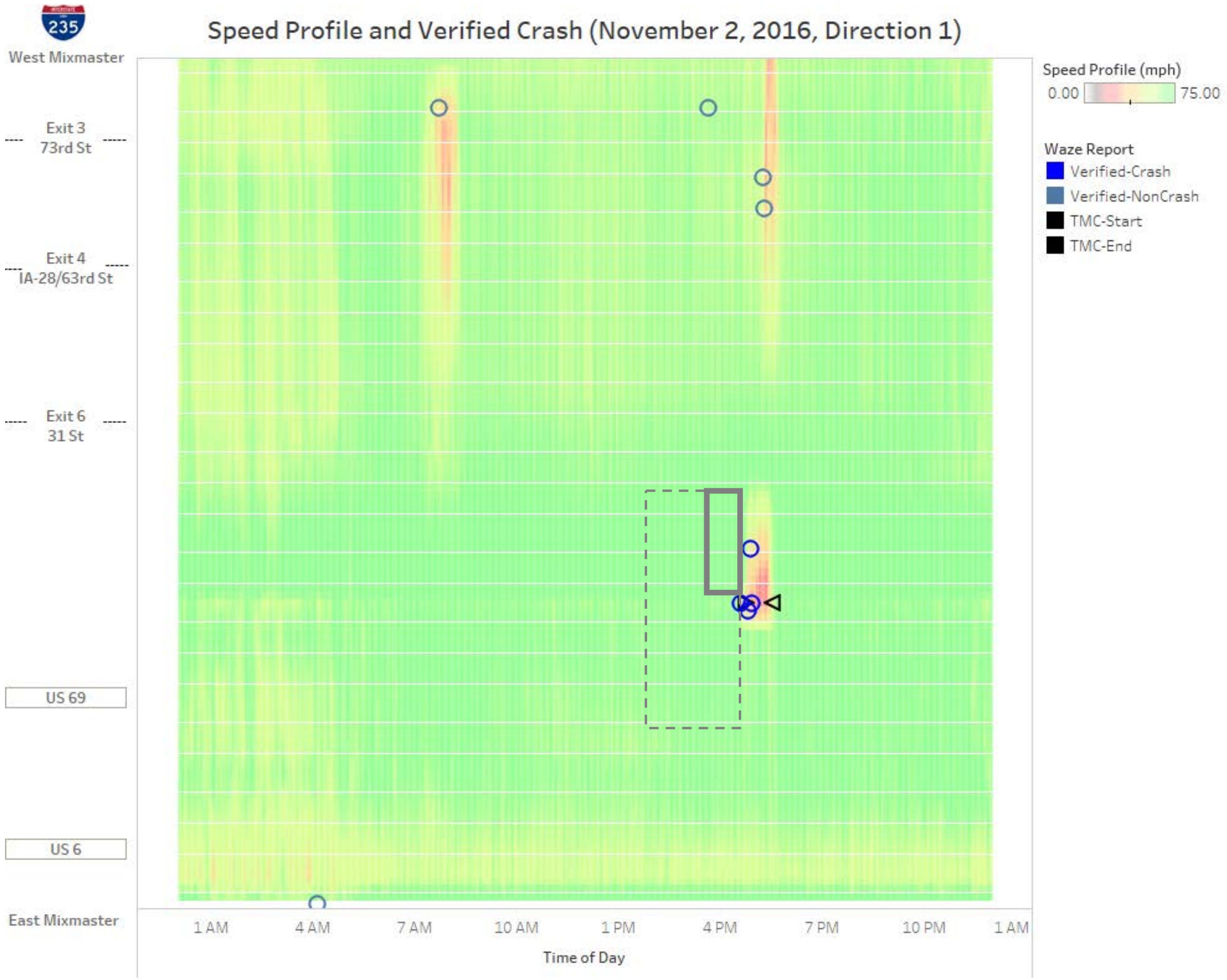
Westbound



Data Preparation – Features

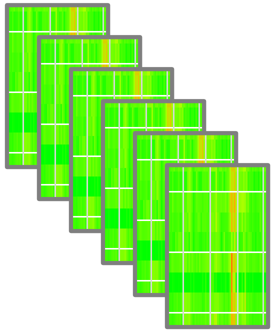
1 mile
(10 pixels)

5 minutes
(5 pixels)

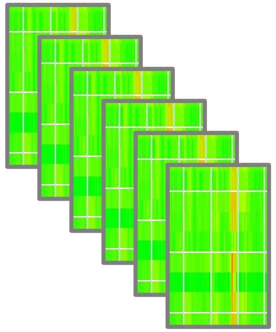


Data Preparation – Summary

- 462 out of 2,840 Waze Reports are labeled as Crash
- 73,014 samples Generated
- 9 Features Included
- Balanced Sample: <u>7,920 Crashes vs. 7,920 Non-Crashes</u>
- 70% train 30% validation



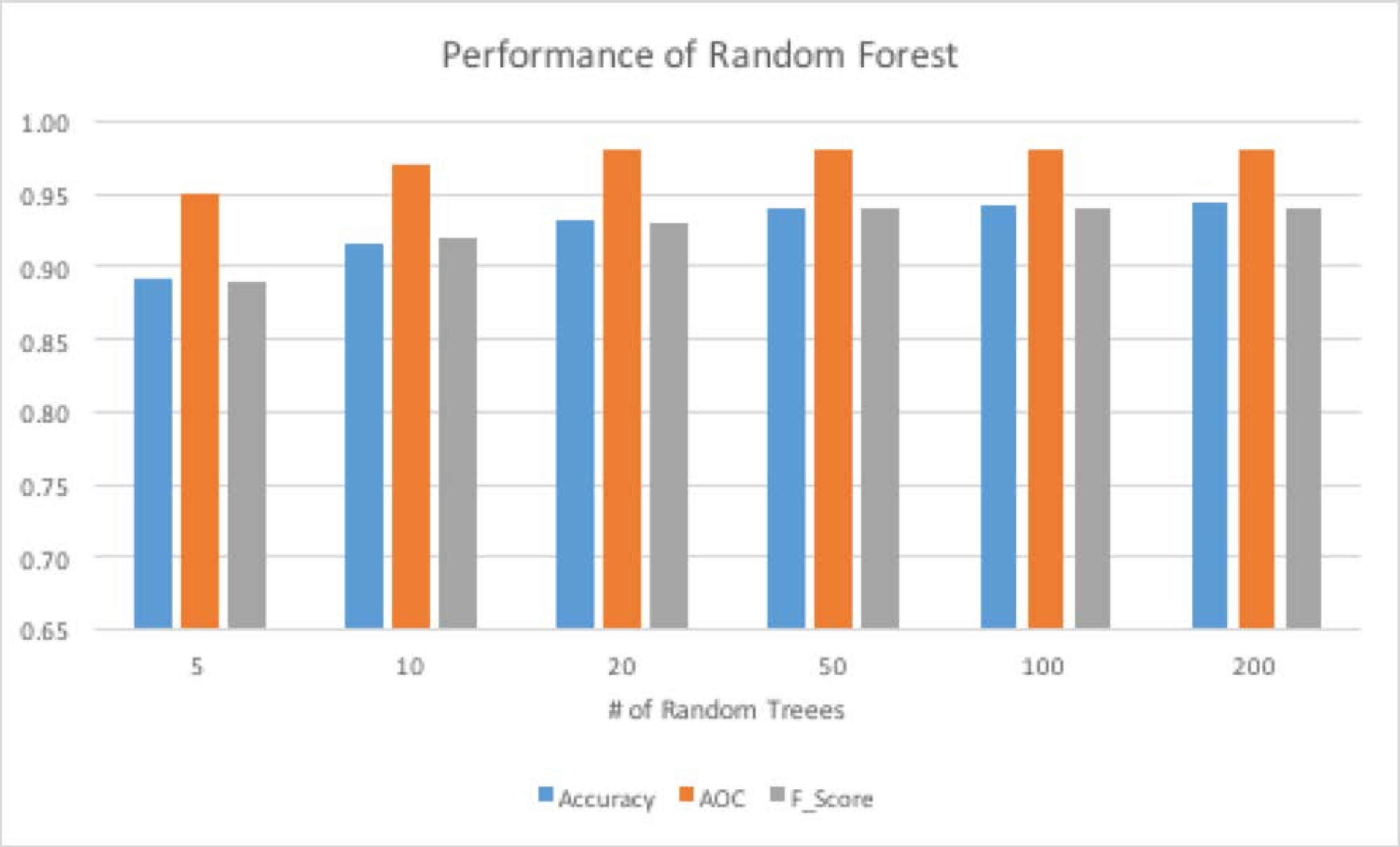
(7920, 10, 5, 9)



(7920, 10, 5, 9)

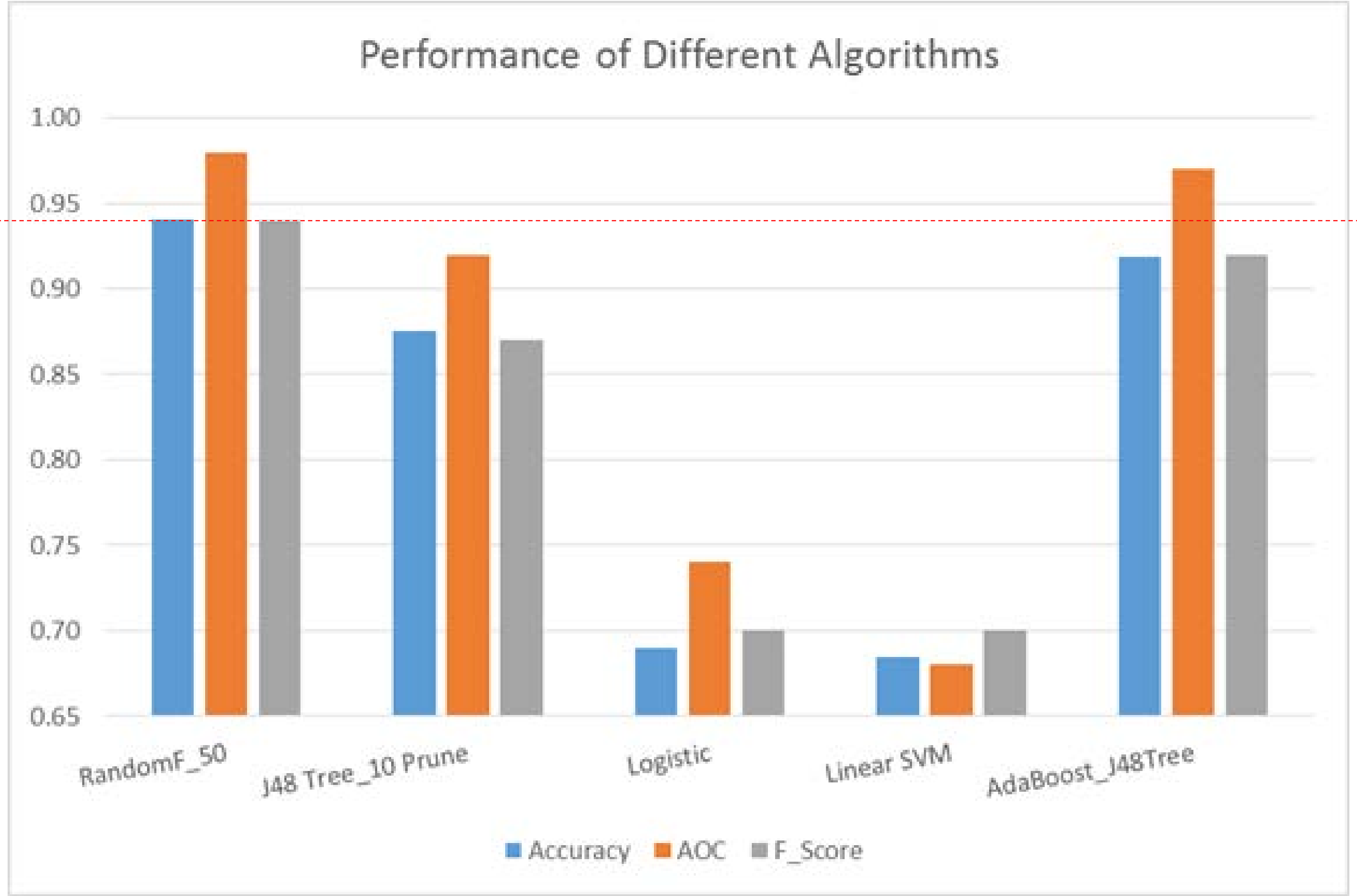
Model Results – Shallow Model

Model Results – Shallow Model



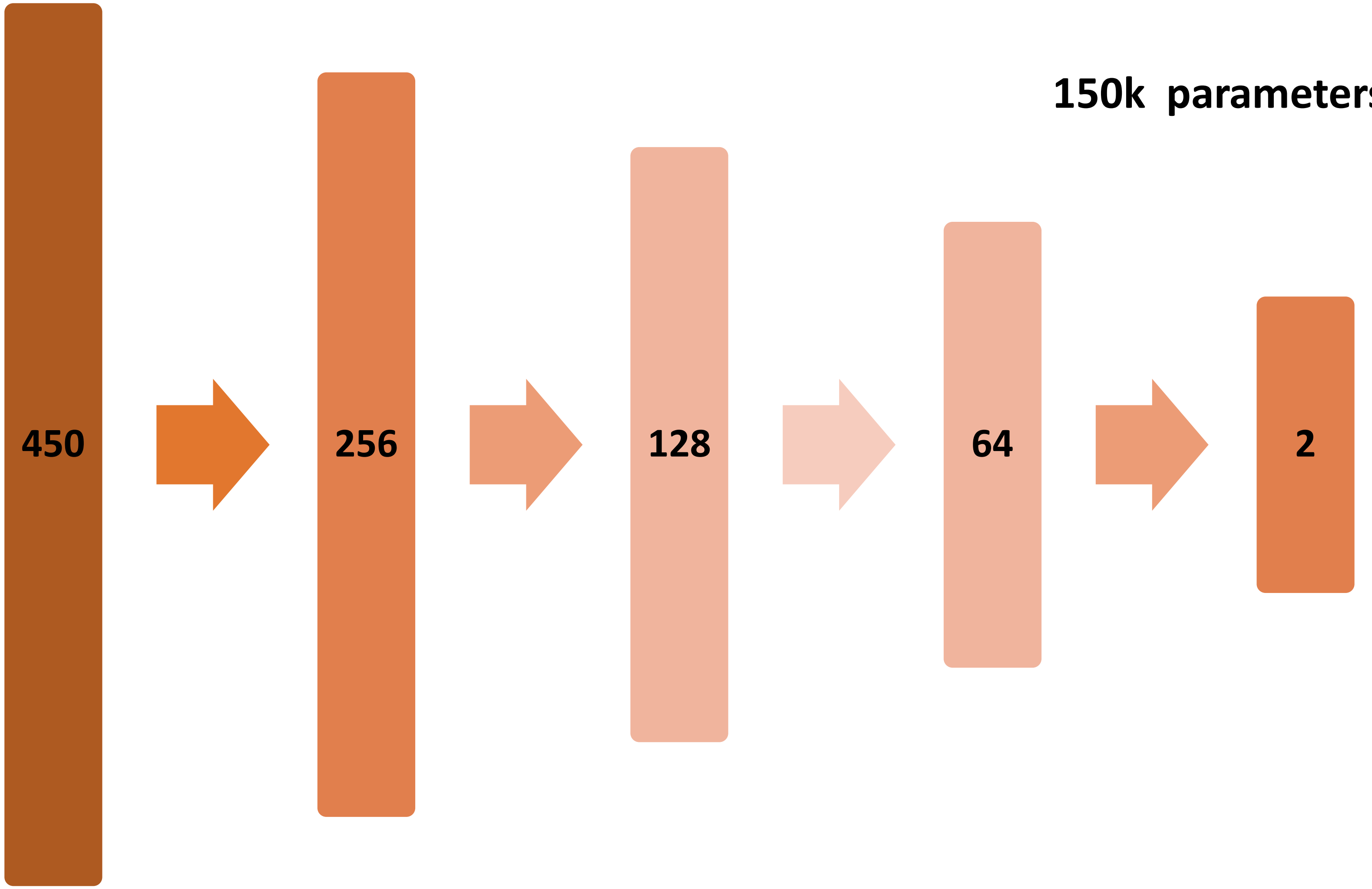
Model Results – Shallow Model

94%



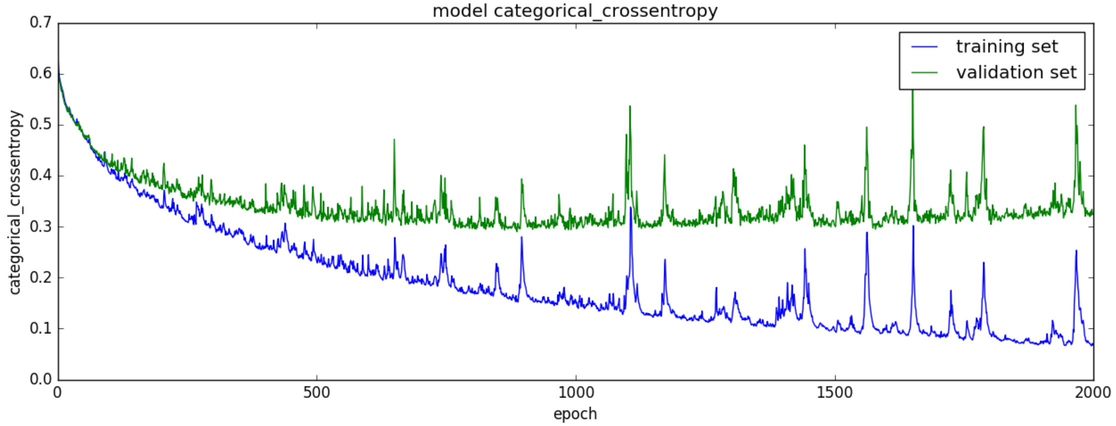
Model Results – Deep Model (FC)

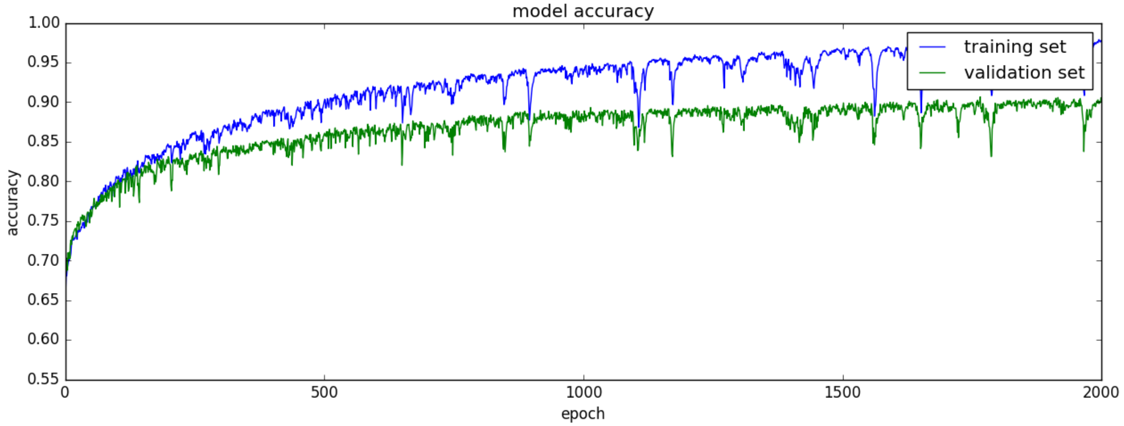
150k parameters



Model Results – Deep Model (FC)

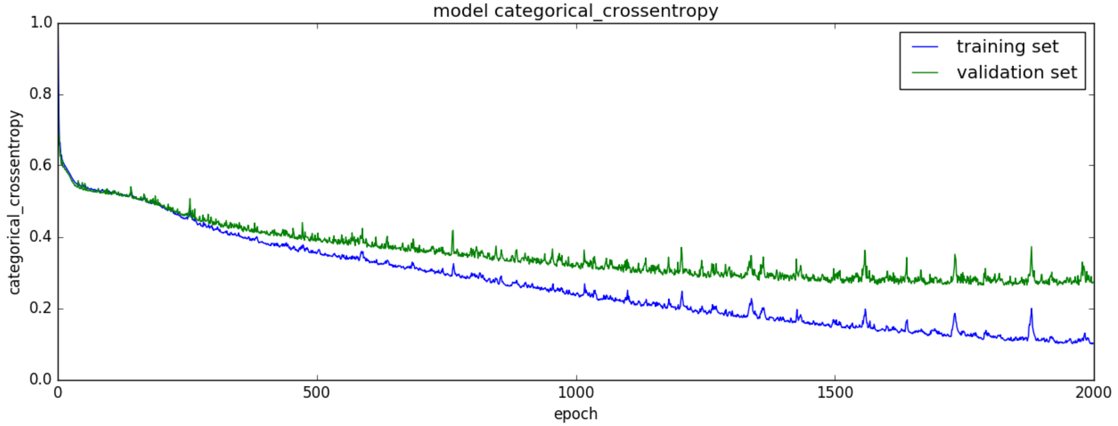
Activation	Epoch	Batch Normalization	Drop Out
ELU	2K	Yes	ALL, 0.25
RELU		No	No

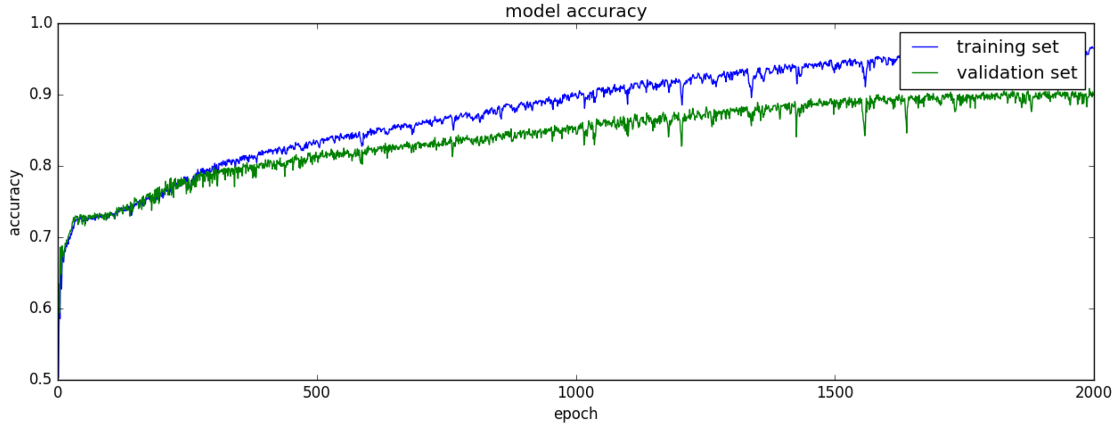




Model Results – Deep Model (FC)

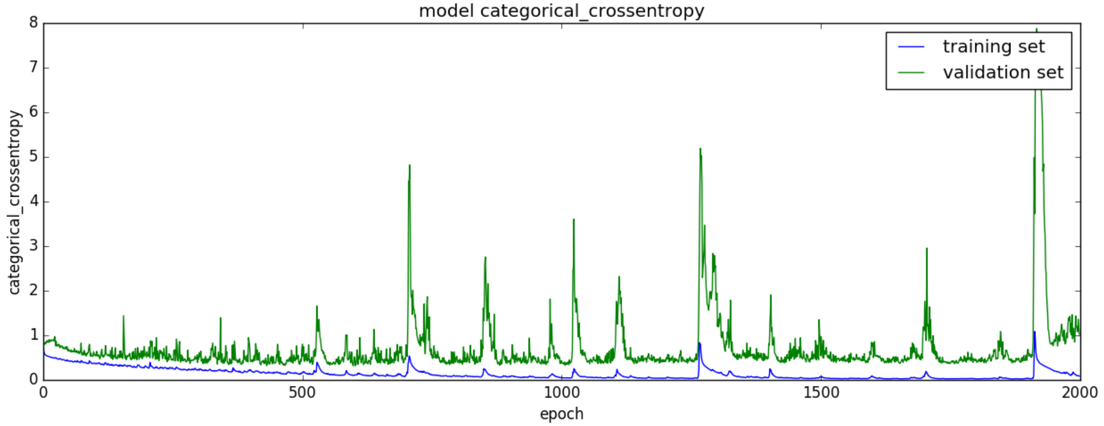
Activation	Epoch	Batch Normalization	Drop Out
ELU	2K	Yes	ALL, 0.25
RELU		No	No

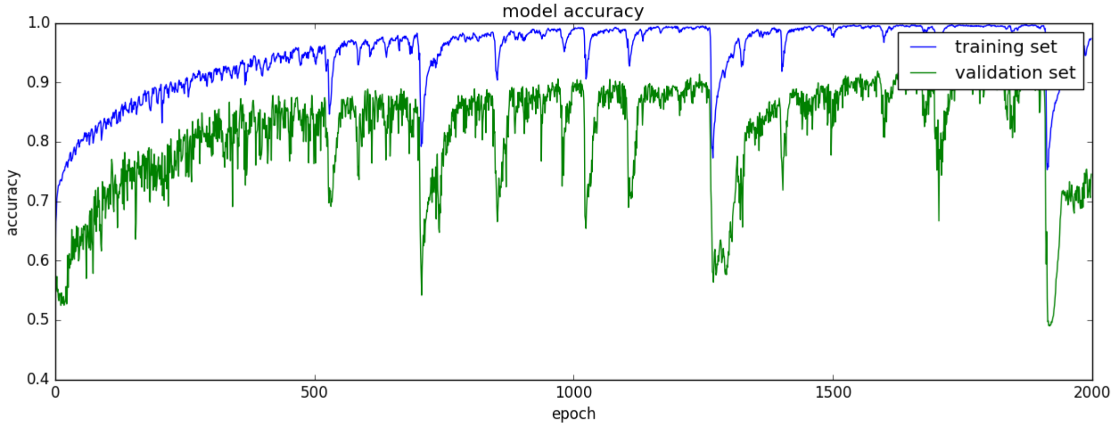




Model Results – Deep Model (FC)

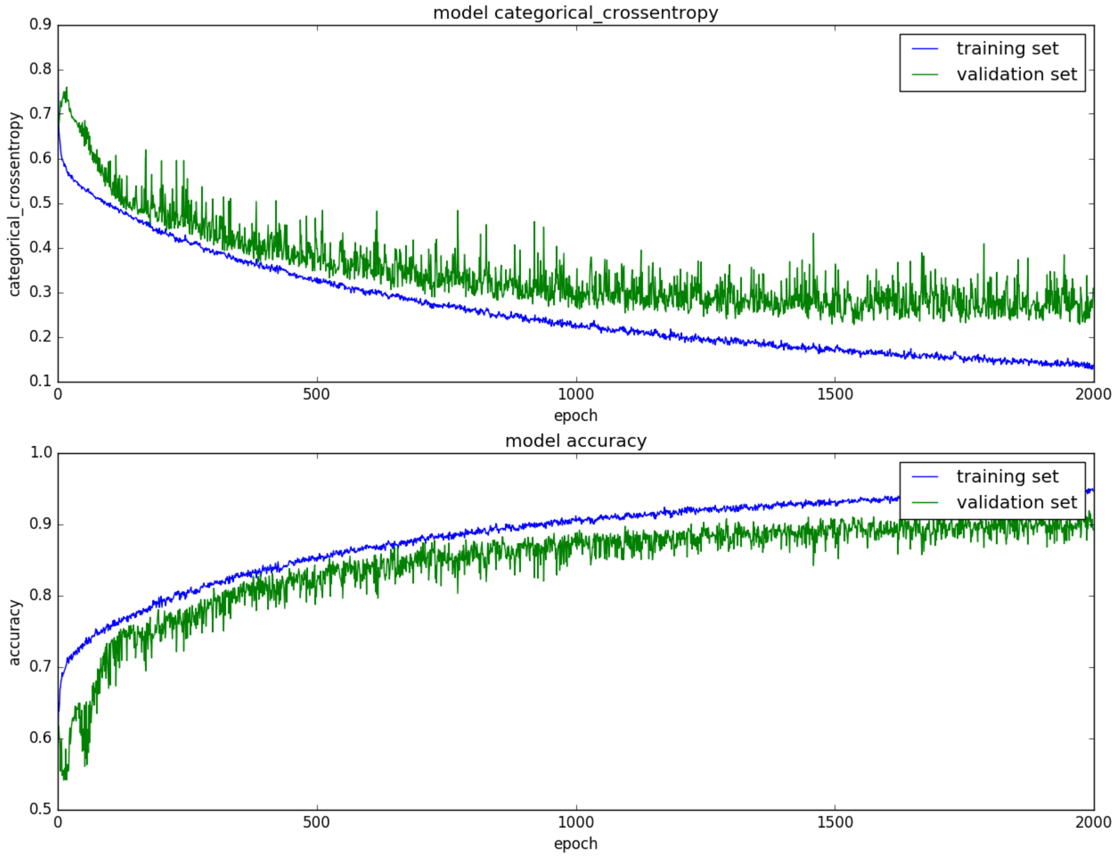
Activation	Epoch	Batch Normalization	Drop Out
ELU	2K	Yes	ALL, 0.25
RELU		No	No





Model Results – Deep Model (FC)

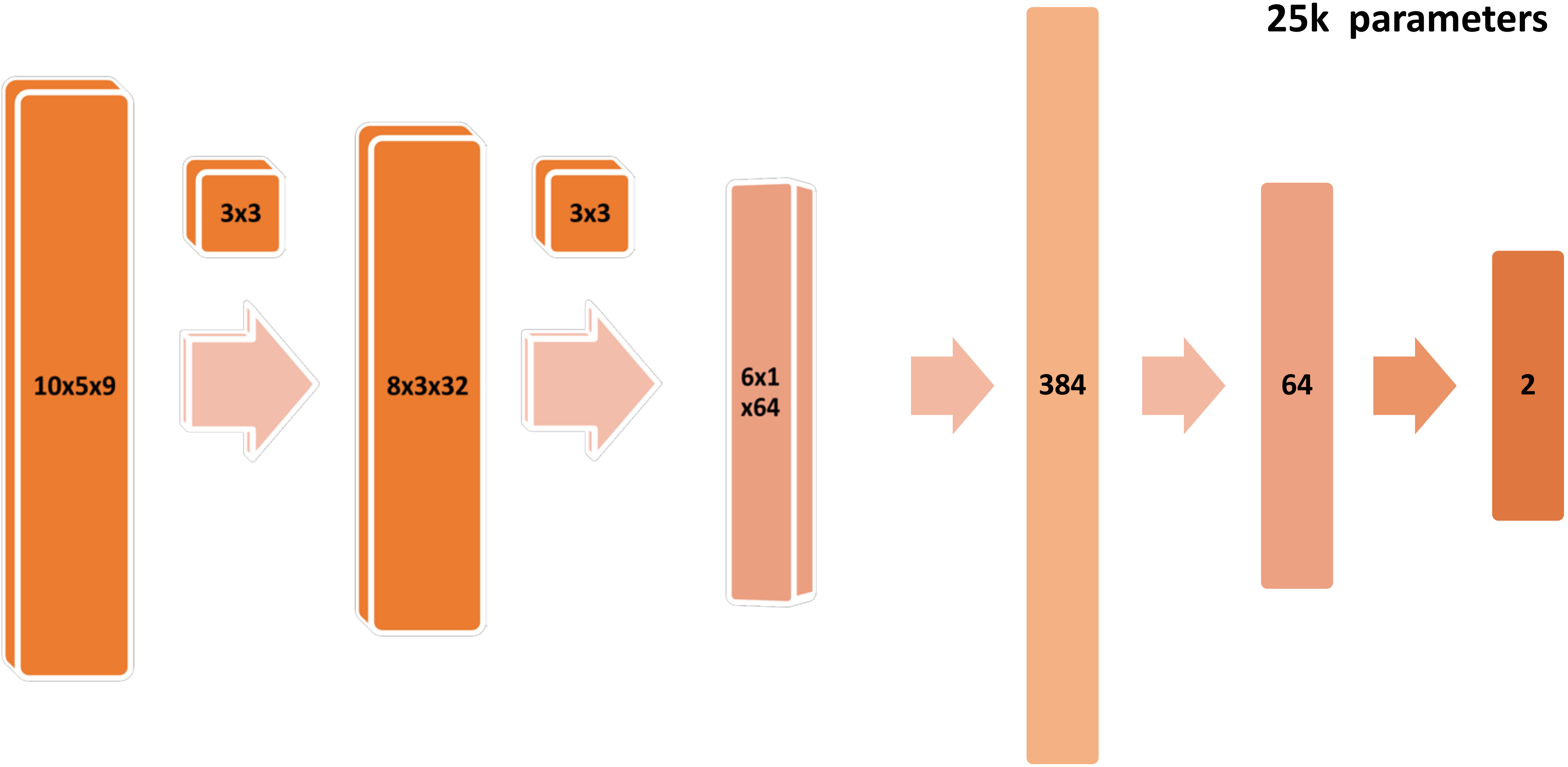
Activation	Epoch	Batch Normalization	Drop Out
ELU	2K	Yes	ALL, 0.25
RELU		No	No



90%

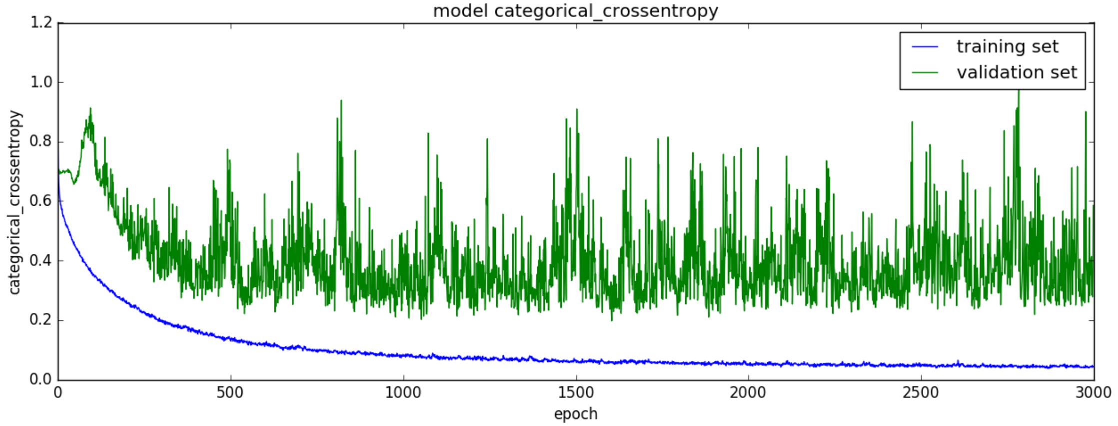
Model Results – Deep Model (CNN)

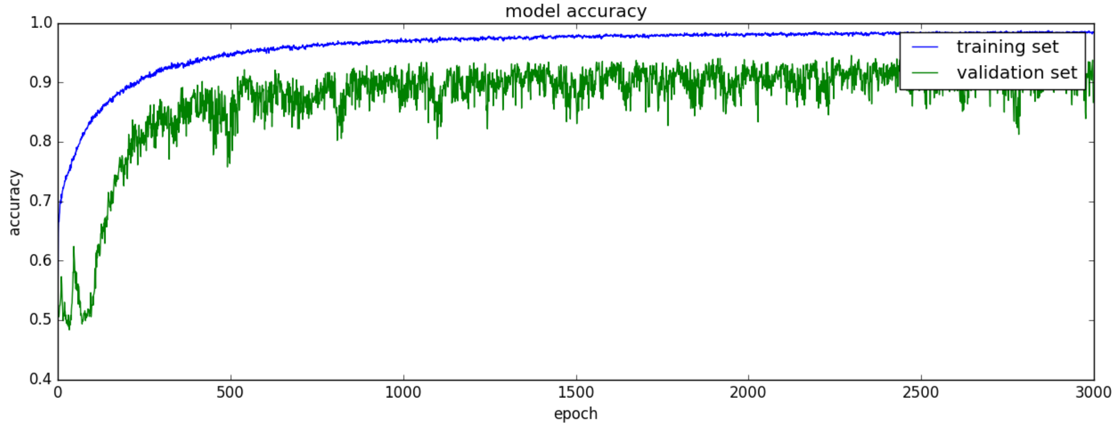
25k parameters



Model Results – Deep Model (CNN)

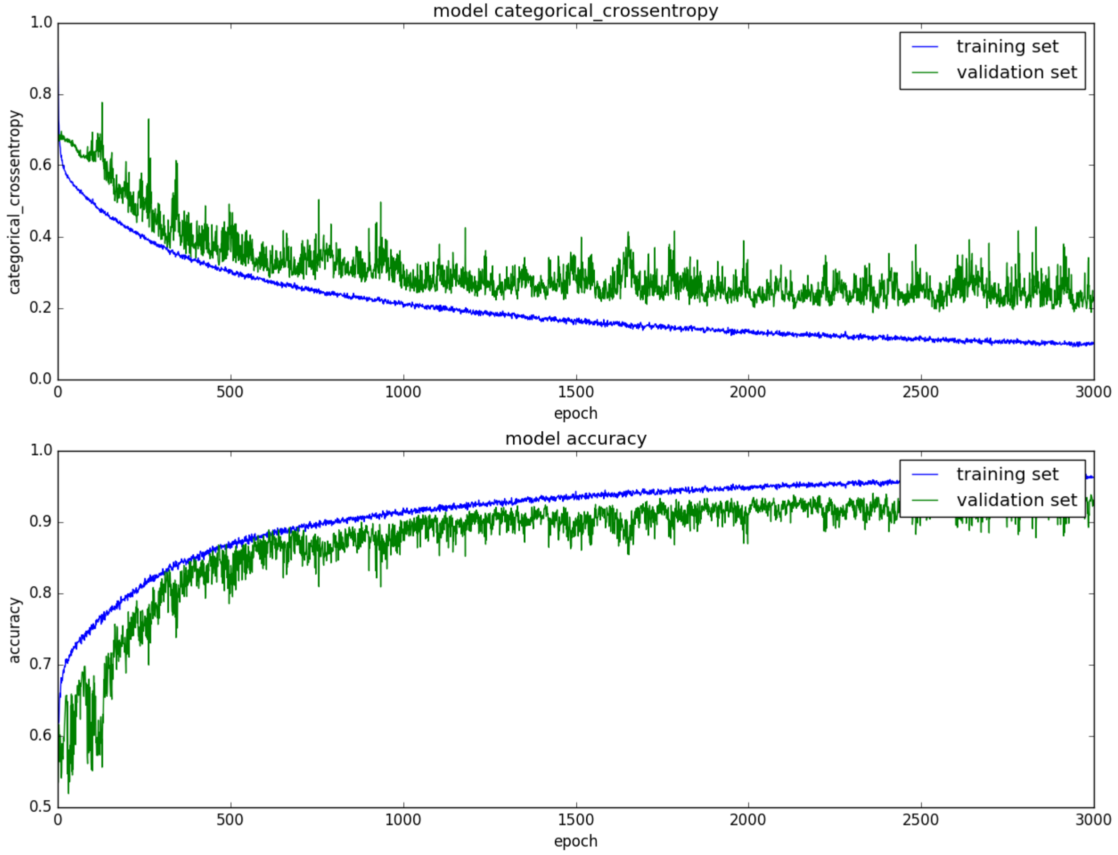
Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5





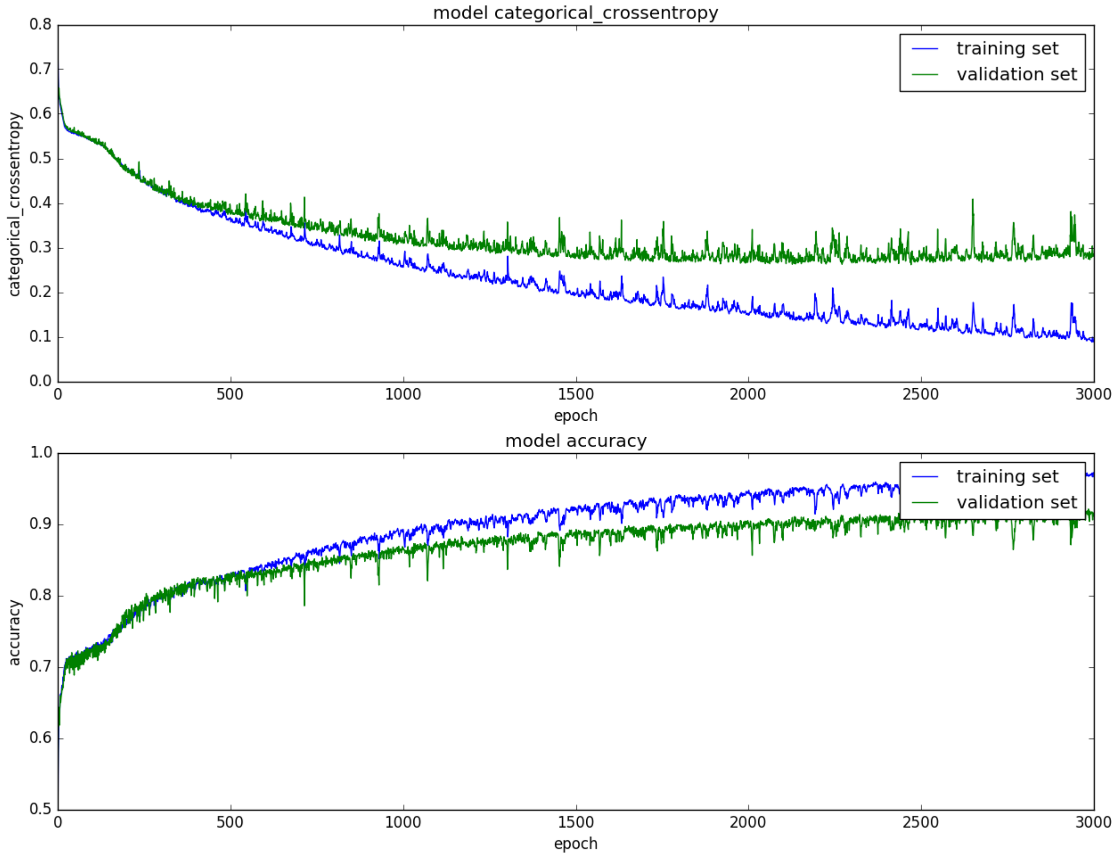
Model Results – Deep Model (CNN)

Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



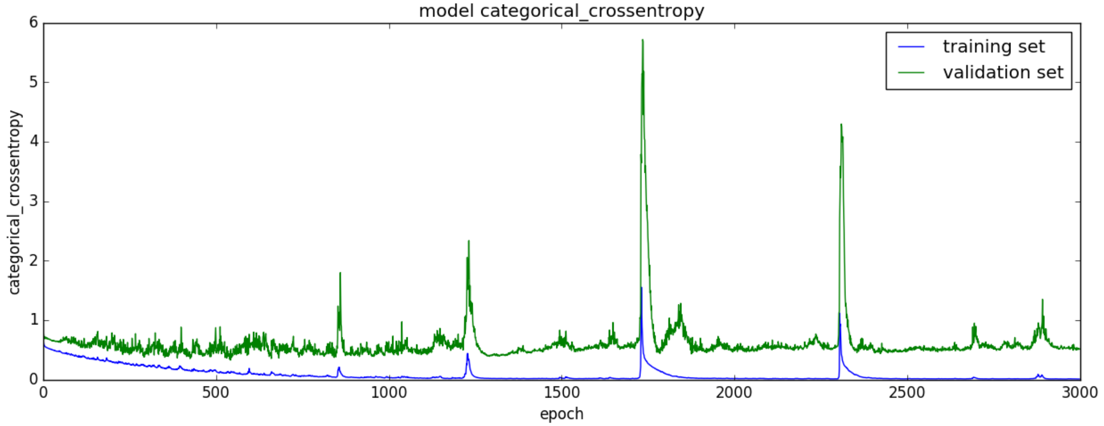
Model Results – Deep Model (CNN)

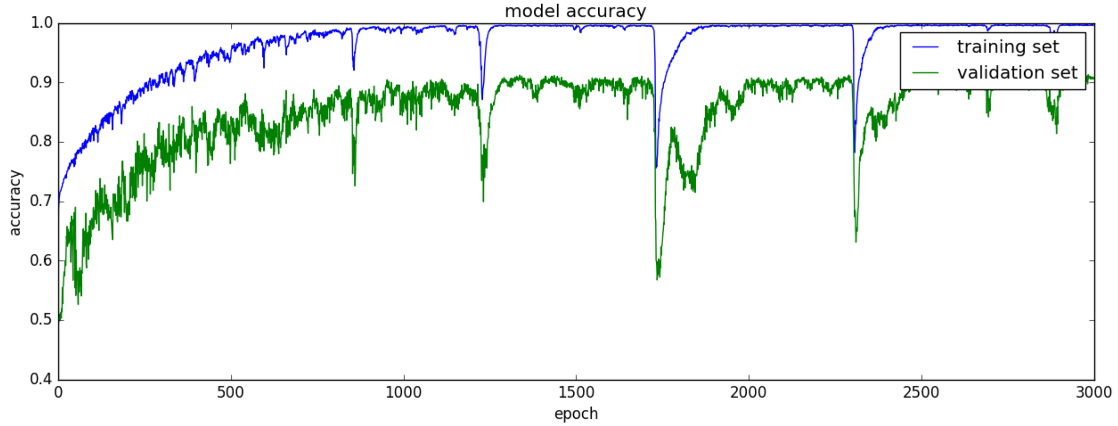
Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



Model Results – Deep Model (CNN)

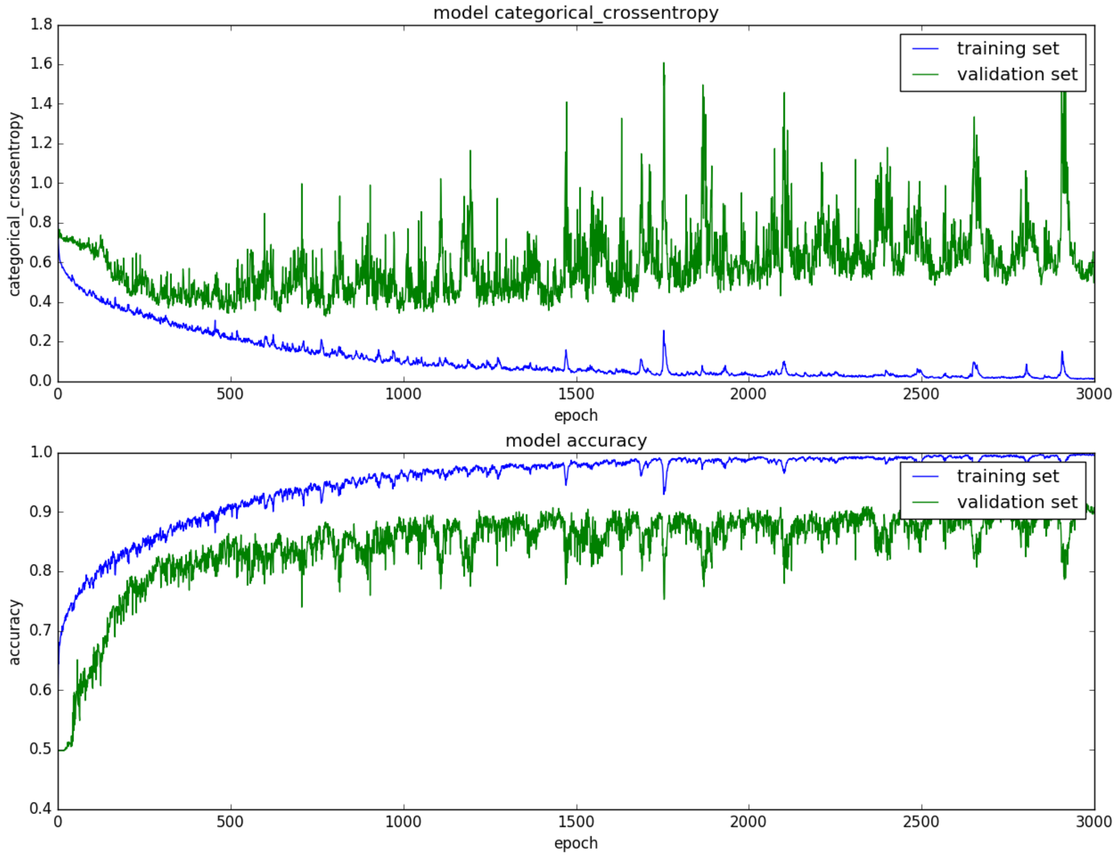
Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5





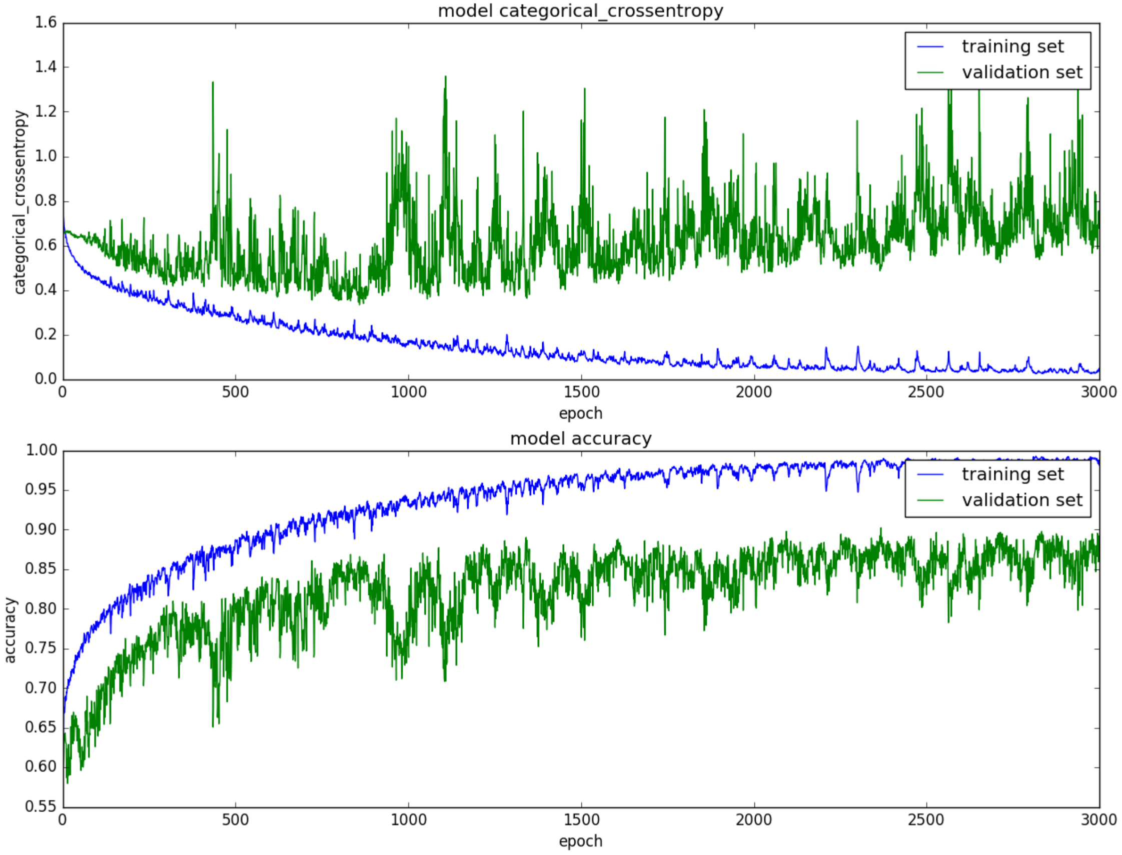
Model Results – Deep Model (CNN)

Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



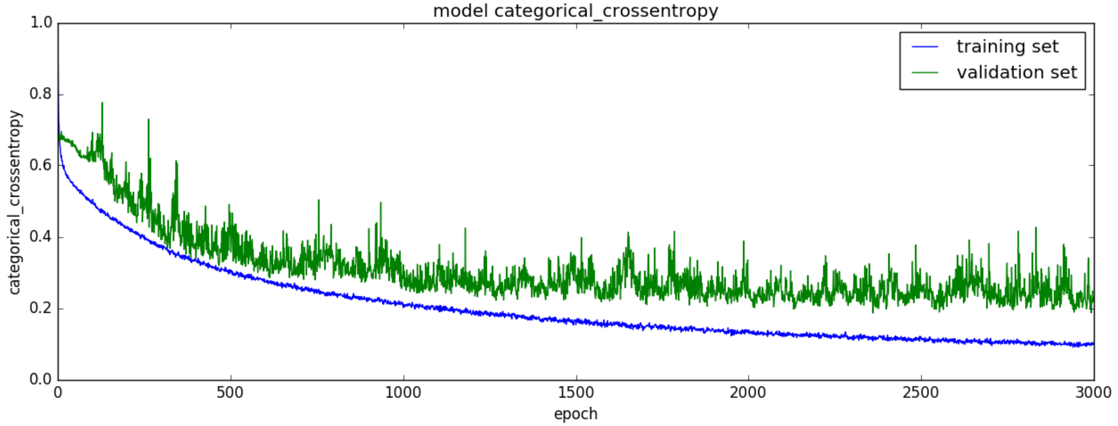
Model Results – Deep Model (CNN)

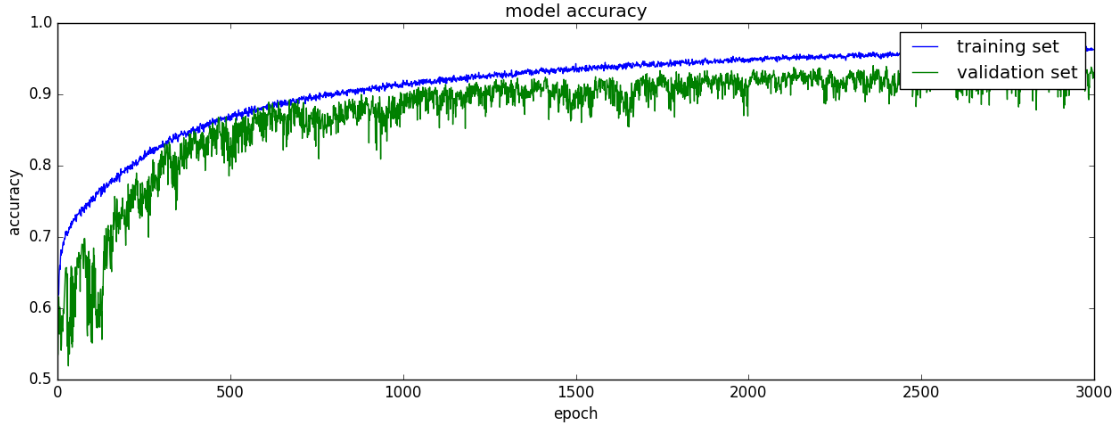
Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



Model Results – Deep Model (CNN)

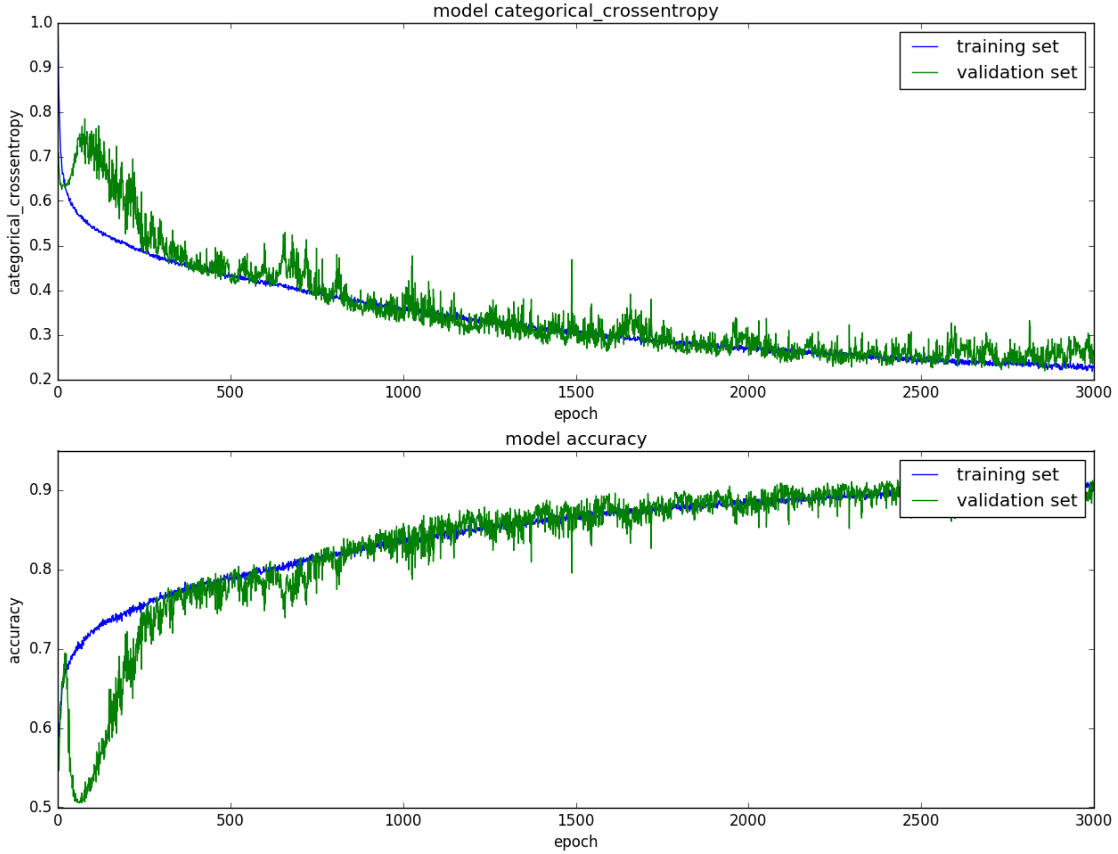
Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5





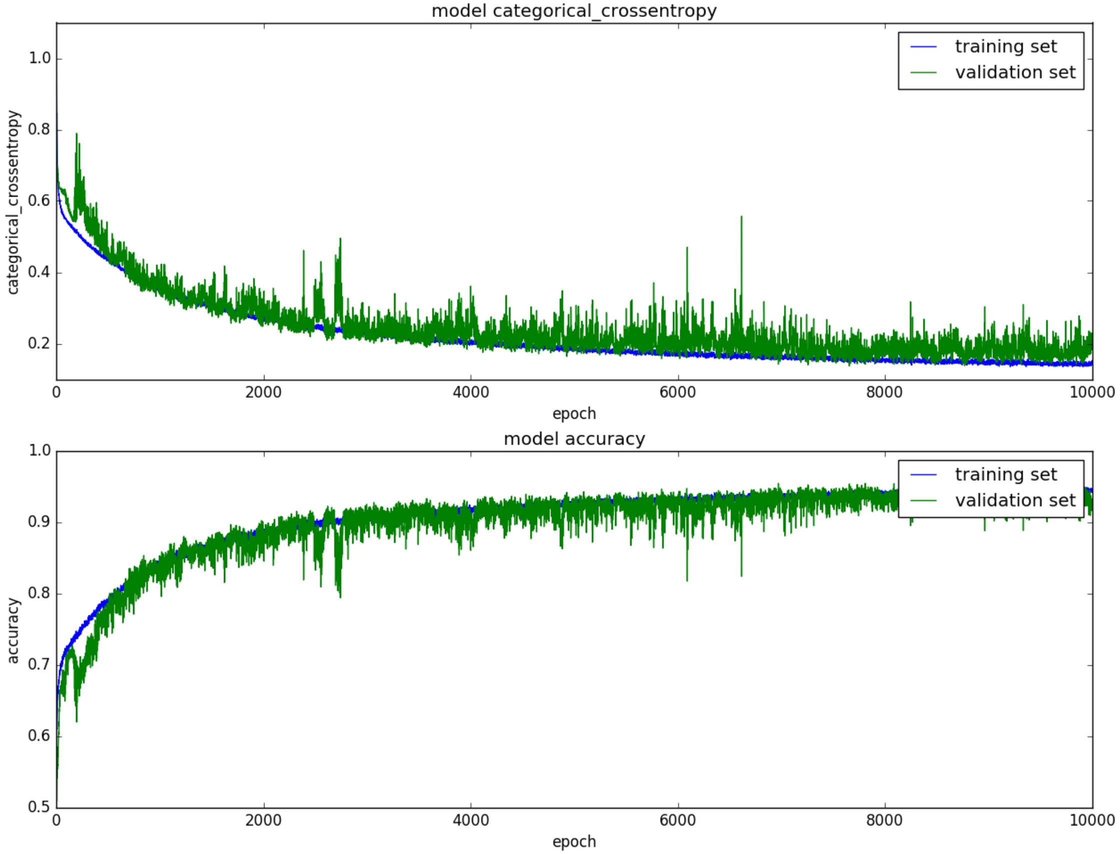
Model Results – Deep Model (CNN)

Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



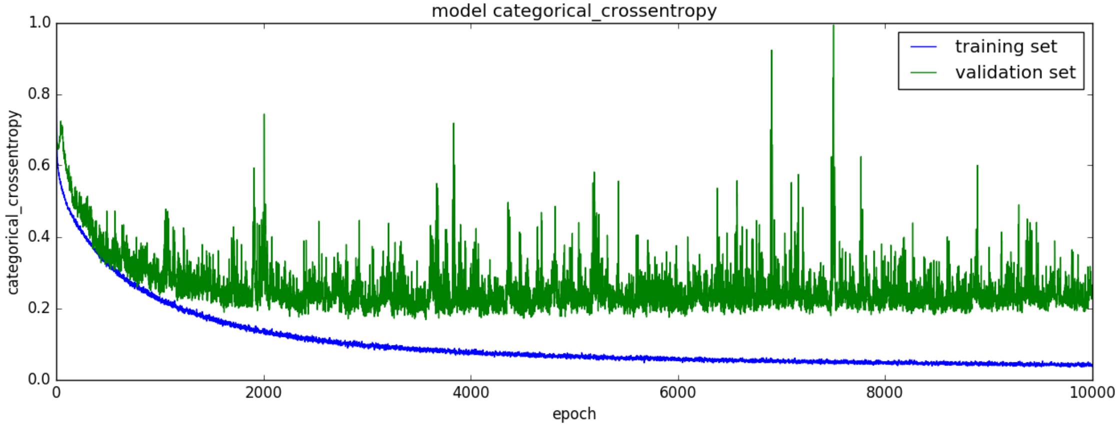
Model Results – Deep Model (CNN)

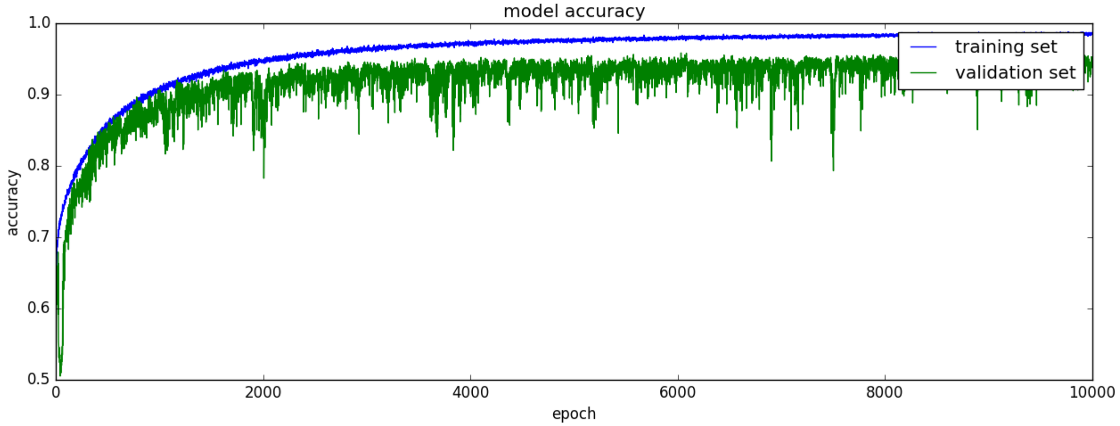
Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



Model Results – Deep Model (Best)

Activation	Epoch	Batch Normalization	Drop Out	Drop Out Position	Drop Out Rate
ELU	3K	Yes	Yes	ALL	0.25
RELU	10K	No	No	LAST	0.5



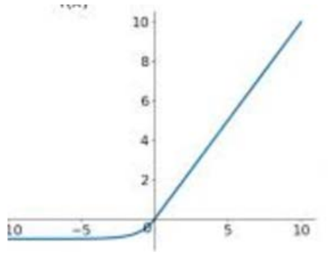


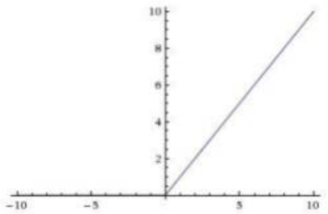
Train： 99.3%

Validation: 95.54%

Discussion

- Random forest can compete with deep models while shallow models can not
- ELU is more stable than ReLU in term of training





- Batch normalization helps training accuracy a lot
- Dropout to all hidden layer can better reduce overfitting
- CNN is better than FC
- Best accuracy: Train 99.3% Validation: 95.54%