Friedrich-Alexander-Universität
Technische Fakultät

Machine Learning
Data Analytics

FAU

Predicting the Severity of Support Tickets from Machine Event Data

Shriya Mittapalli

Agenda

Introduction

Motivation

- Difficult to determine the severity of issues
- Delayed ticket resolution and unnecessary on-site visits
- Sub-optimal resource utilization
- Customer dissatisfaction

Introduction

Problem Statement

- Determine whether a problem can be resolved remotely or requires on-site intervention
- Address the challenge of unreliable severity predictions
- Research and implement uncertainty estimations in deep neural networks

Introduction
Goal

- Research solutions for uncertainty estimations in deep neural networks
- Model provides understandable insights into predictions

Introduction

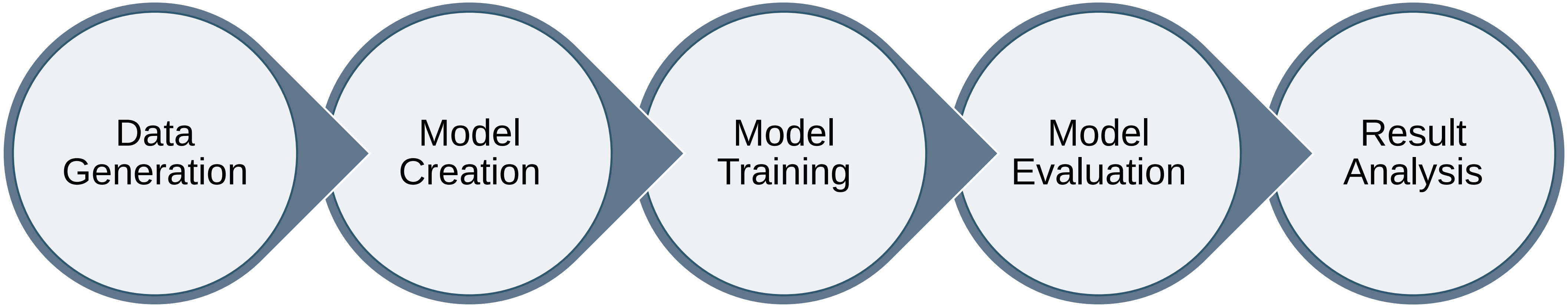
Related Work

Evaluating pointwise reliability of machine learning prediction.
[Giovanna Nicora et al., 2022, Journal of Biomedical Informatics]

Can You Trust This Prediction? Auditing Pointwise Reliability After Learning.
[Peter Schulam et al., 2019, arXiv:1901.00403]

A Survey of Uncertainty in Deep Neural Networks
[Jakob Gawlikowski et al., 2022, arXiv:2107.03342]

Methodology



Data Generation
Dataset

CT Scan machine - Siemens Healthineers

The dataset includes:

- ❑Case ID (Ticket ID)
- ❑Machine ID
- ❑Onsite(Class Label)
- ❑Start and End time of the ticket
- ❑Sequence of Events

Case_Id	Machine_Id	Onsite	Start_Time	End_Time
0	Machine_0	1	2014-12-21 10:54:31+00:00	2014-12-23 23:14:34+00:00
1	Machine_1	1	2016-05-05 07:21:48+00:00	2016-09-23 04:47:57+00:00
2	Machine_2	0	2016-07-17 12:38:41+00:00	2016-07-19 04:13:16+00:00
3	Machine_1	0	2016-07-19 13:23:57+00:00	2016-08-05 03:13:37+00:00

Event	Timestamp	Case_Id
126	2016-07-17 12:30:20+00:00	2
328	2016-07-17 12:26:30+00:00	2
206	2016-07-17 12:24:42+00:00	2

Data Generation

Data Preprocessing

Combine events files

↓

Identify events to discarded

↓

Identify Case IDs to be discarded

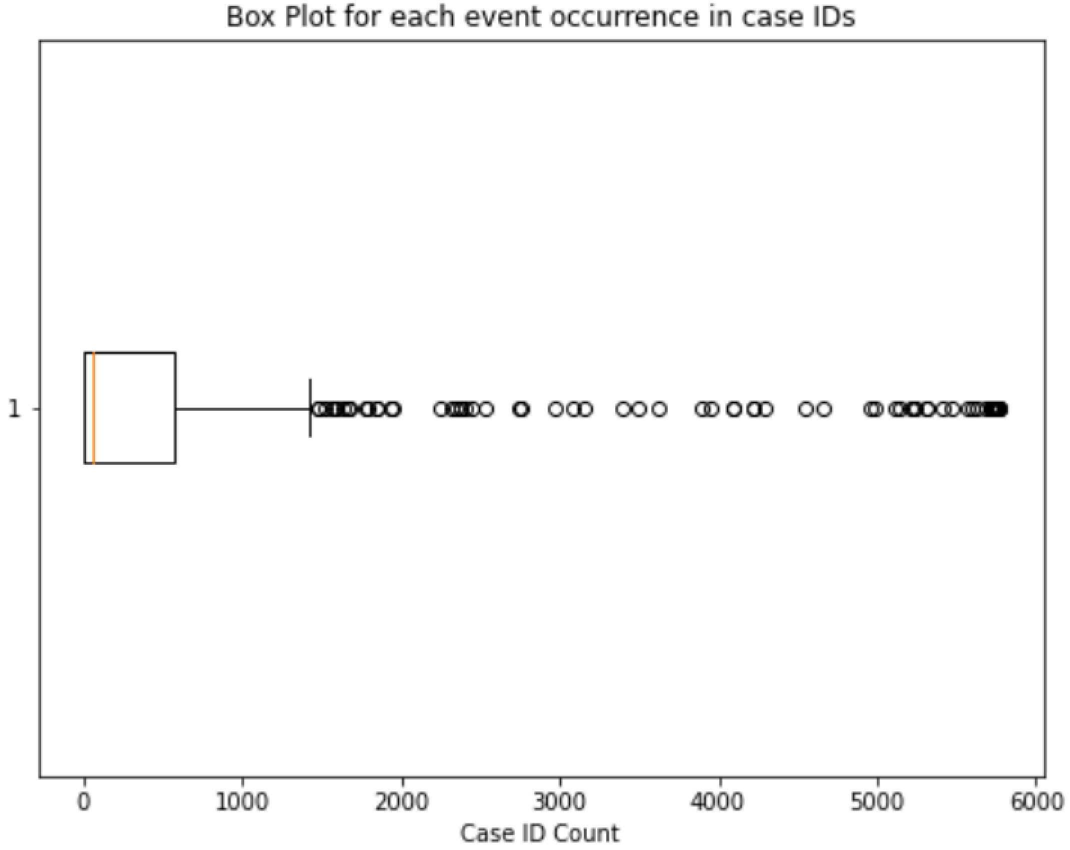
↓

Get last 8000 events and create final dataset

Data Generation

Data Preprocessing

	Event	Case_Id_Count
0	382	5709
1	303	580
2	220	4288
3	201	4099
4	206	5763

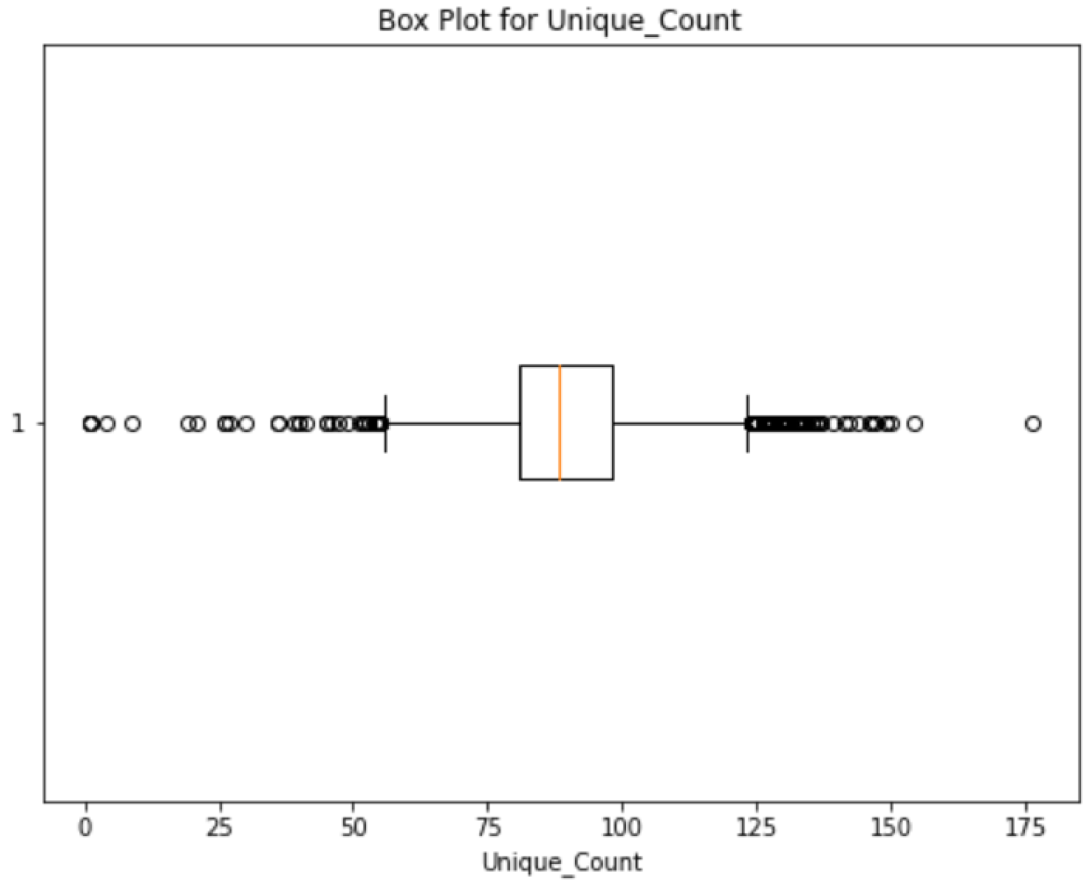


❖ 285 events removed

Data Generation

Data Preprocessing

	Case_Id	Unique_Count
0	11	91
1	205	90
2	246	91
3	261	99
4	315	74



❖ 10 Case IDs removed

Data Generation

Final Data

	Case_Id	Last_8000_Events	Unique_Count	Machine_Id	Onsite	Start_Time	End_Time
0	0	66,304,66,304,66,304,572,597,328,206,434,206,3...	82	Machine_0	1.0	2014-12-21 10:54:31+00:00	2014-12-23 23:14:34+00:00
1	1	328,206,206,328,328,328,562,572,572,572,572,26...	75	Machine_1	1.0	2016-05-05 07:21:48+00:00	2016-09-23 04:47:57+00:00
2	2	328,206,434,206,414,328,492,528,516,206,55,64,...	81	Machine_2	1.0	2016-07-17 12:38:41+00:00	2016-07-19 04:13:16+00:00
3	4	64,166,527,119,317,317,240,298,599,129,280,260...	100	Machine_4	1.0	2015-11-07 08:25:12+00:00	2015-11-14 02:24:16+00:00
5	6	64,166,122,25,381,560,35,122,25,381,560,35,230...	96	Machine_6	1.0	2016-01-24 03:00:55+00:00	2016-01-31 23:23:50+00:00

Model Creation

Uncertainty Models

1. Monte Carlo Dropout Model

2. Ensemble Model
   a. 4 Dense Model with different initial weights
   b. Gated Recurrent Unit (GRU)

Weights:

- ❑ Glorot Normal
- ❑ He Normal
- ❑ Orthogonal
- ❑ Variance Scaling

Model Creation

Monte Carlo Dropout Model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 flatten (Flatten)           (None, 8000)              0

 dense (Dense)               (None, 8192)              65544192

 dropout (Dropout)           (None, 8192)              0

 dense_1 (Dense)             (None, 4096)              33558528

 dropout_1 (Dropout)         (None, 4096)              0

 dense_2 (Dense)             (None, 2048)              8390656

 dropout_2 (Dropout)         (None, 2048)              0

 dense_3 (Dense)             (None, 1024)              2098176

 dropout_3 (Dropout)         (None, 1024)              0

 dense_4 (Dense)             (None, 512)               524800

 dropout_4 (Dropout)         (None, 512)               0

 dense_5 (Dense)             (None, 256)               131328

 dropout_5 (Dropout)         (None, 256)               0

 dense_6 (Dense)             (None, 128)               32896

 dropout_6 (Dropout)         (None, 128)               0

 dense_7 (Dense)             (None, 64)                8256

 dropout_7 (Dropout)         (None, 64)                0

 dense_8 (Dense)             (None, 32)                2080

 dropout_8 (Dropout)         (None, 32)                0

 dense_9 (Dense)             (None, 1)                 33

=================================================================
Total params: 110290945 (420.73 MB)
Trainable params: 110290945 (420.73 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Model Creation
Ensemble Model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 flatten (Flatten)           (None, 8000)              0

 dense (Dense)               (None, 8192)              65544192

 dense_1 (Dense)             (None, 4096)              33558528

 dense_2 (Dense)             (None, 2048)              8390656

 dense_3 (Dense)             (None, 1024)              2098176

 dense_4 (Dense)             (None, 512)               524800

 dense_5 (Dense)             (None, 256)               131328

 dense_6 (Dense)             (None, 126)               32382

 dense_7 (Dense)             (None, 64)                8128

 dense_8 (Dense)             (None, 32)                2080

 dropout (Dropout)           (None, 32)                0

 dense_9 (Dense)             (None, 1)                 33

=================================================================
Total params: 110290303 (420.72 MB)
Trainable params: 110290303 (420.72 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Dense Model

Model Creation

Ensemble Model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 flatten (Flatten)           (None, 8000)              0

 dense (Dense)               (None, 8192)              65544192

 dense_1 (Dense)             (None, 4096)              33558528

 dense_2 (Dense)             (None, 2048)              8390656

 dense_3 (Dense)             (None, 1024)              2098176

 dense_4 (Dense)             (None, 512)               524800

 dense_5 (Dense)             (None, 256)               131328

 dense_6 (Dense)             (None, 126)               32382

 dense_7 (Dense)             (None, 64)                8128

 dense_8 (Dense)             (None, 32)                2080

 dropout (Dropout)           (None, 32)                0

 dense_9 (Dense)             (None, 1)                 33

=================================================================
Total params: 110290303 (420.72 MB)
Trainable params: 110290303 (420.72 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Dense Model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 gru (GRU)                   (None, 8000, 128)         50304

 dropout (Dropout)           (None, 8000, 128)         0

 gru_1 (GRU)                 (None, 64)                37248

 dropout_1 (Dropout)         (None, 64)                0

 dense (Dense)               (None, 1)                 65

=================================================================
Total params: 87617 (342.25 KB)
Trainable params: 87617 (342.25 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

GRU Model

Model Training

Data split: train, validation, and test (60%,20%,20%)

Optimizer: Adam

Metrics: Uncertainty, Accuracy

Loss: Binary Cross Entropy

EarlyStopping: monitor - validation loss, min delta - 0.001, patience - 5

Model Evaluation

Approach 1: Probability Range

- Average prediction probabilities
- Identify Threshold range
- Calculate uncertainty

Model Evaluation

Approach 1: Probability Range

- Average prediction probabilities
- Identify Threshold range
- Calculate uncertainty

Approach 2: Standard Deviation

- Average prediction probabilities
- Standard Deviation of model
- Average probability $\pm$ Standard Deviation
- Calculate uncertainty

Model Evaluation

$$\text{Uncertainity} = \frac{\text{Uncertain Predictions}}{Total\ Samples}$$

$$Accuracy = \frac{Correct\ Predictions}{Total\ Predictions}$$

Correct Predictions – number of samples model predicted accurately

Uncertain Predictions – number of samples not classified by model

Total Samples – Total number of samples

Total Predictions – Total Samples - Uncertain Predictions

Result Analysis

Model	Measure	Validation	Test
Monte Carlo Dropout Model	Uncertainty	74.53	74.45
	Accuracy	50.45	55.66
Ensemble Model	Uncertainty	90.04	89.31
	Accuracy	59.14	56.44

Probability Range

Result Analysis

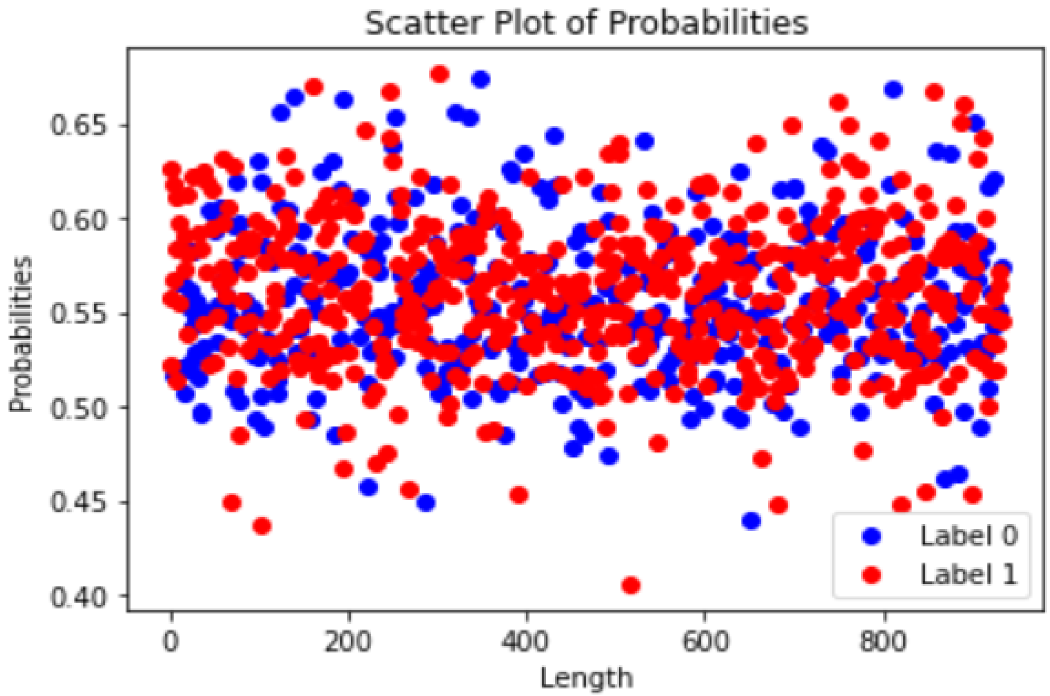
Model	Measure	Validation	Test	SD
Monte Carlo Dropout Model	Uncertainty	93.05	92.02	0.13
	Accuracy	50.00	50.72	
Ensemble Model	Uncertainty	57.67	57.67	0.0104
	Accuracy	50.00	52.50	

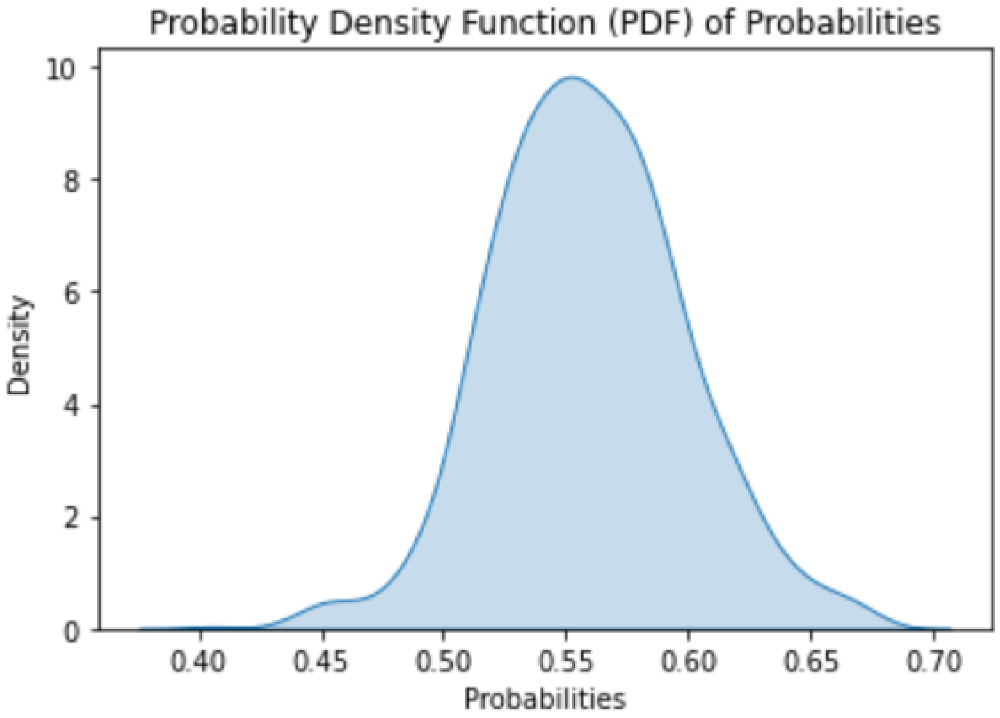
Standard Deviation

Result Analysis

Monte Carlo Dropout Model

Validation Data

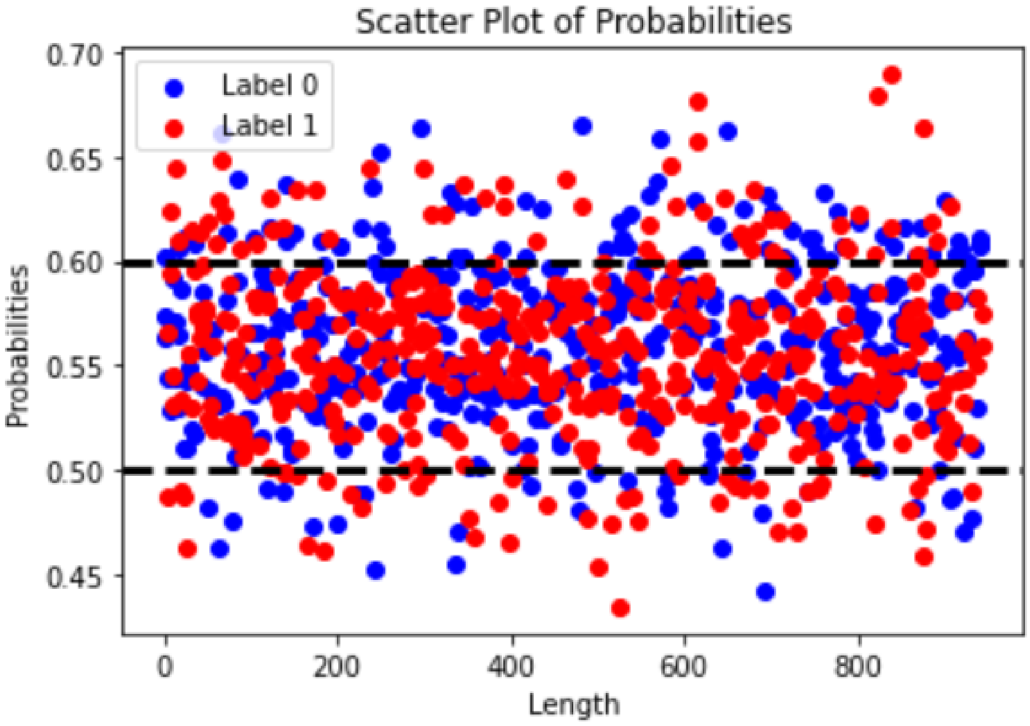


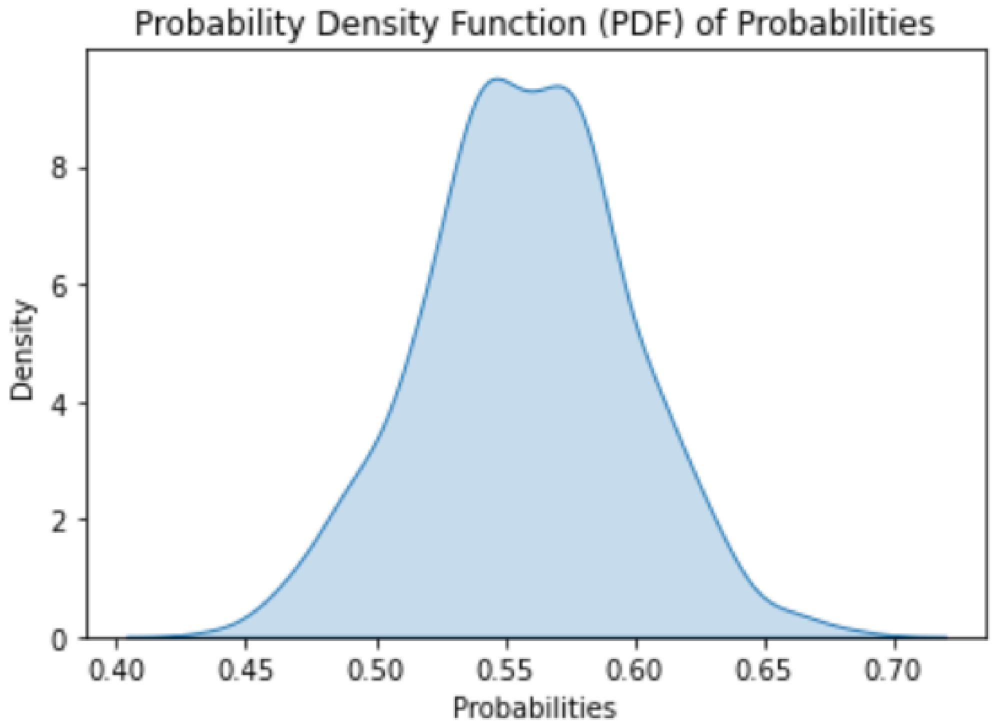


Result Analysis

Monte Carlo Dropout Model

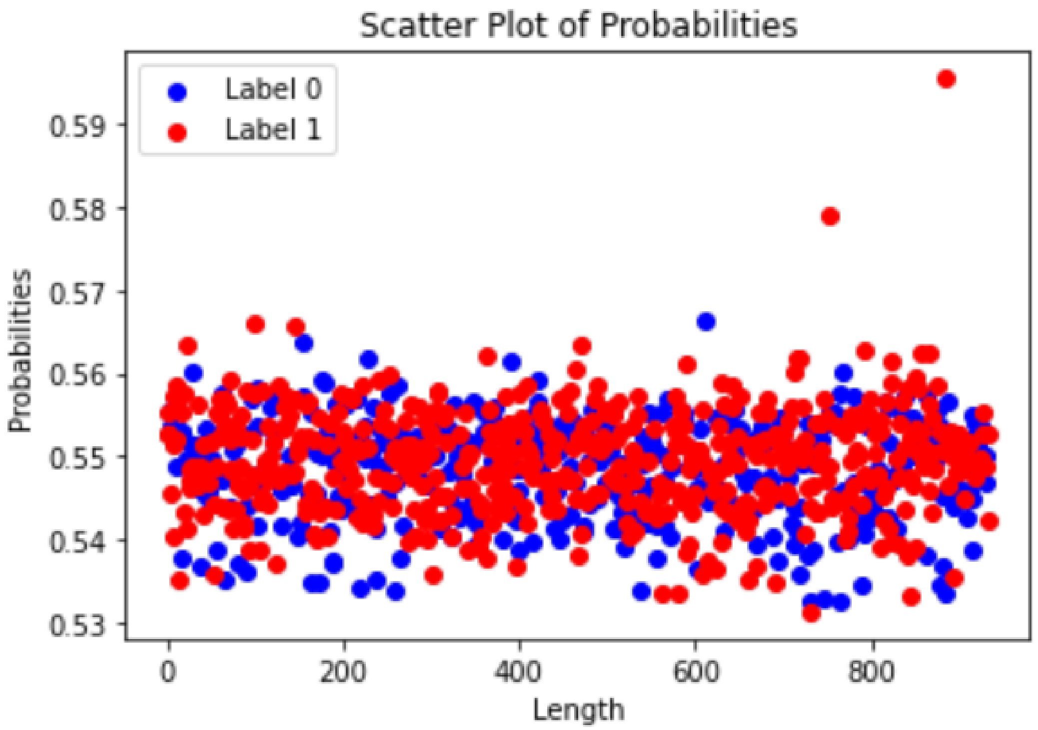
Test Data

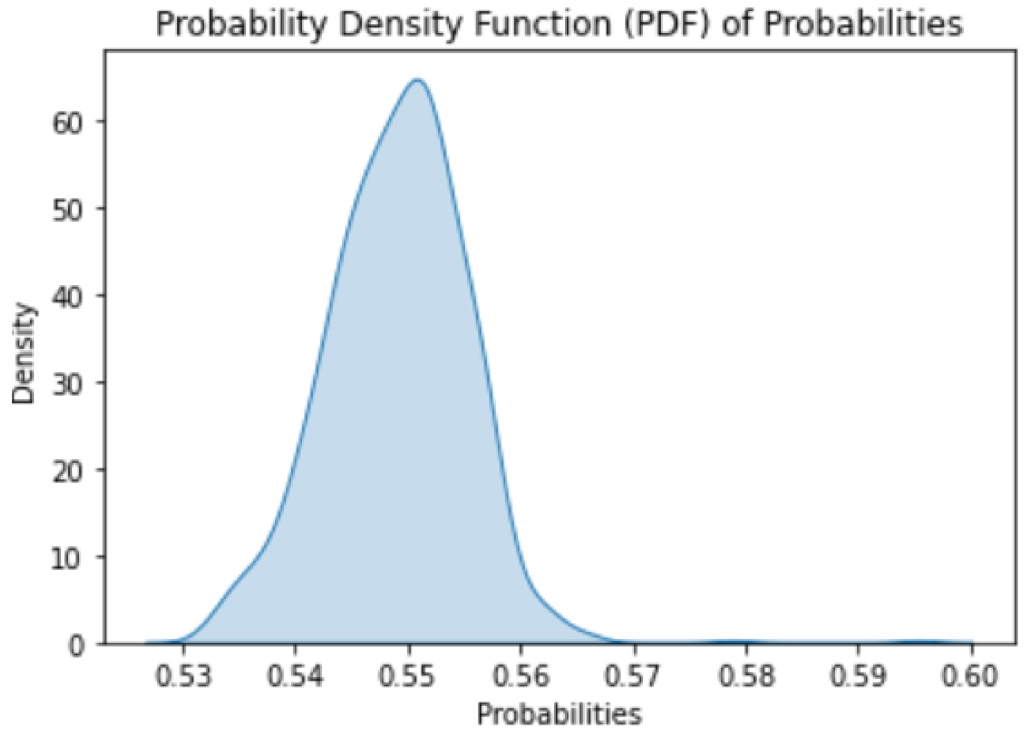




Result Analysis
Ensemble Model

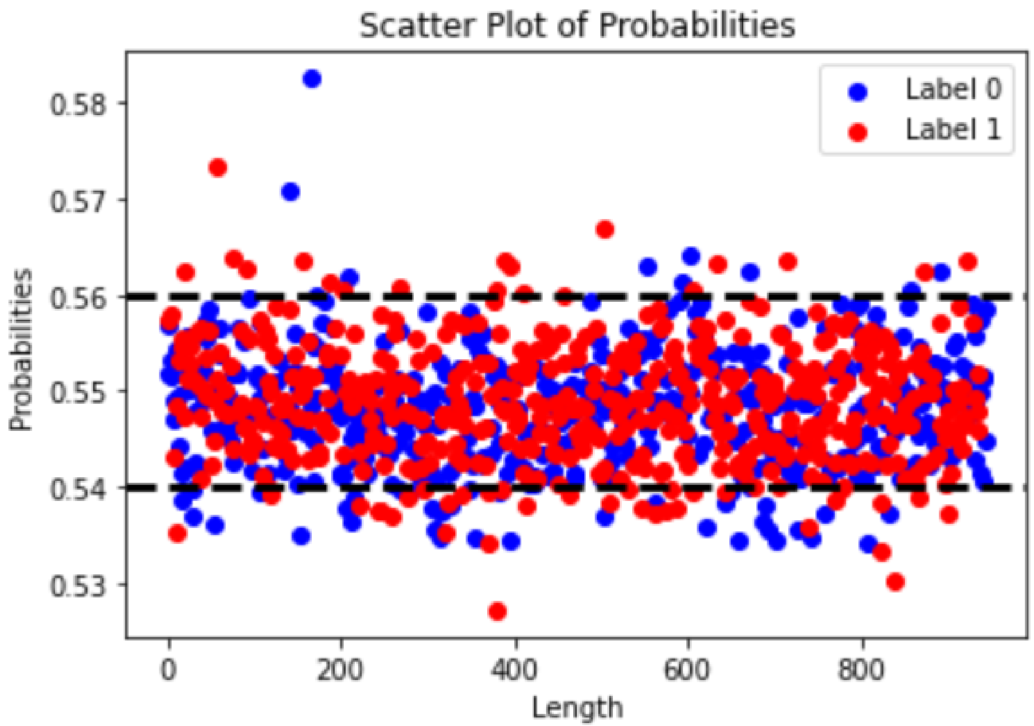
Validation Data

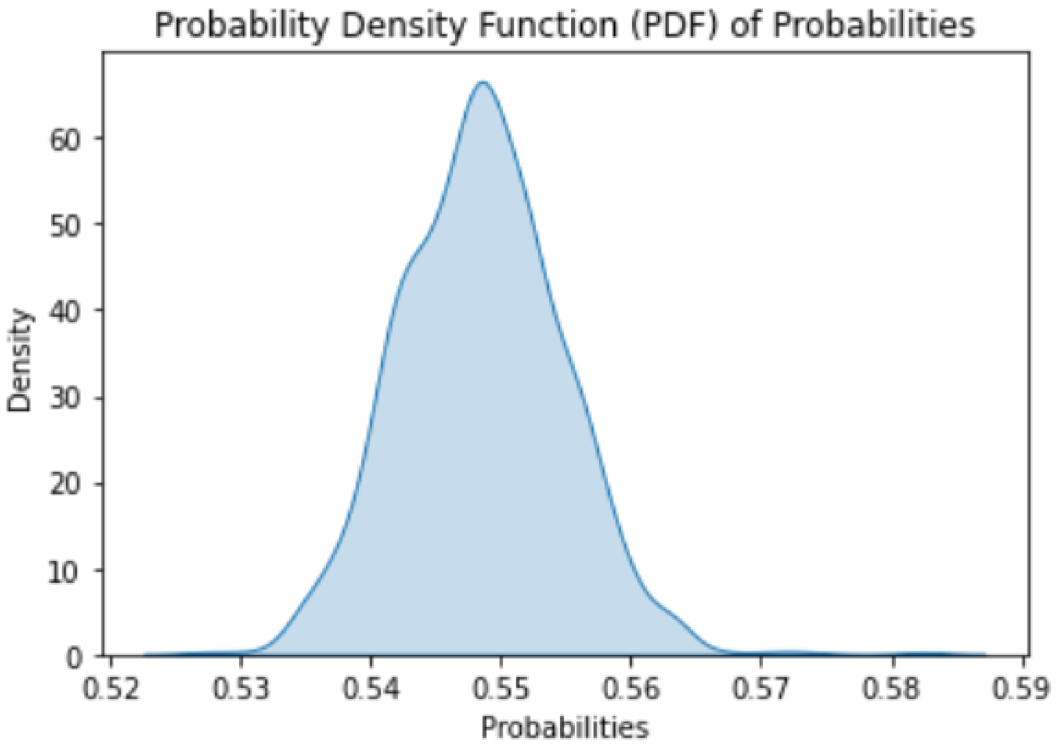




Result Analysis
Ensemble Model

Test Data





Conclusion

- Classify Tickets
- Provide reliability measure

Considerations:

- Standard Deviation – different models
- Threshold boundaries – model uncertainty
- Data limitation constraints
- Transformer Models and LSTM with Gaussian Mixture Models (LSTM-GMMS)


ONE DOES NOT SIMPLY
HAVE ZERO UNCERTAINTY
imgflip.com

Appendix

Weights

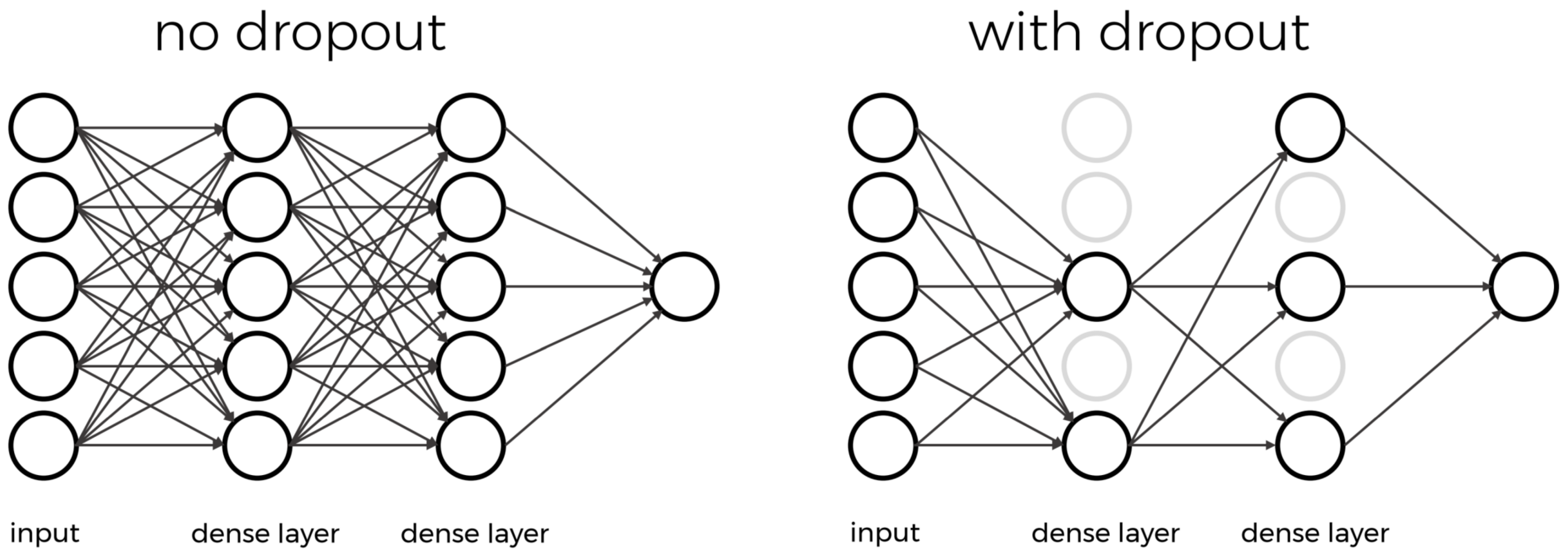
Glorot Normal – Also known as Xavier normal. Draws samples from a truncated normal distribution centered on 0 with stddev = sqrt(2 / (fan_in + fan_out)) where fan_in is the number of input units in the weight tensor and fan_out is the number of output units in the weight tensor.

He Normal – Draws samples from a truncated normal distribution centered on 0 with stddev = sqrt(2 / fan_in) where fan_in is the number of input units in the weight tensor.

Orthogonal – Generates an orthogonal matrix. If the shape of the tensor to initialize is two-dimensional, it is initialized with an orthogonal matrix obtained from the QR decomposition of a matrix of random numbers drawn from a normal distribution.

Variance Scaling – Initializer adapts its scale to the shape of its input tensors. Samples are drawn from a uniform distribution within [-limit, limit], where limit = sqrt(3 * scale / n).

Dense model


no dropout
input
dense layer
dense layer
with dropout
input
dense layer
dense layer

Gated Recurrent Unit (GRU)