SIEMENS Gamesa
RENEWABLE ENERGY

AALBORG UNIVERSITY

Deep Wind: AI Models for Wind Turbine Downtime Predictions (eg. Icing)

Rasmus Dovnborg Frederiksen
Data Scientist & Industrial **PhD-student**
Siemens Gamesa Renewable Energy & Aalborg University

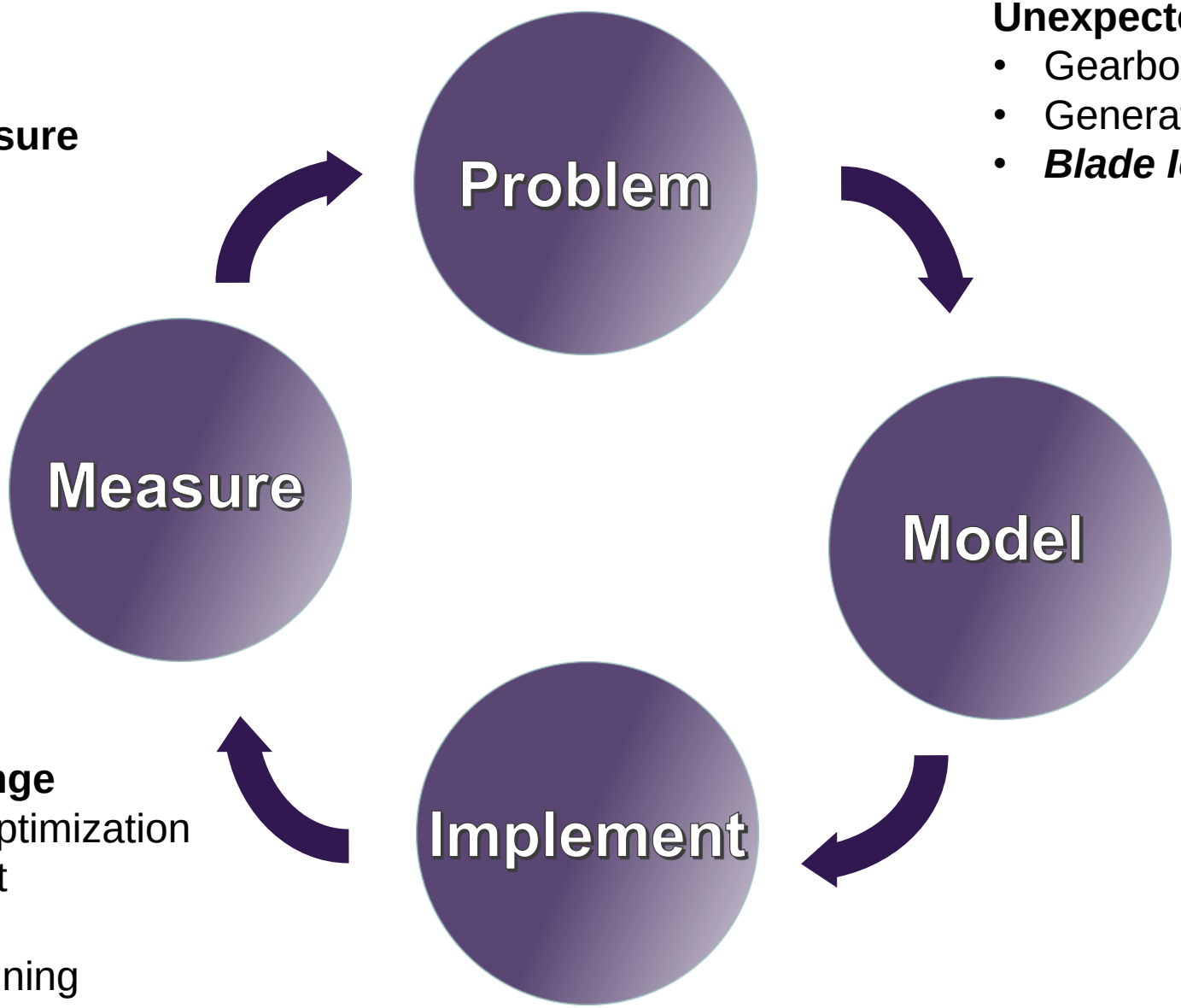
SIEMENS Gamesa
RENEWABLE ENERGY

Agenda:

1. Research
2. Deep Learning
3. Results
4. Caveats and future research

PhD - *Operations Research* – An Applied Science

Unexpected downtime events:
- Gearbox issues
- Generator issues
- ***Blade Icing events!***



Find and measure relevant KPIs
- Cost/Profit
- Throughput
- Waste
- Etc.

Formulate a solution model:
- Simulation
- Statistics
- **Deep Learning**
- Etc.

Implement Change
- Operational Optimization
- AI Deployment
- Redistribution
- Employee Training

Existing Research

Many publications already achieved functional **Deep Learning!**

Literature on ML methods for WTGs					
Reference	ML Method	Data Source	Prediction Target	Purpose	Prediction Horizon
[3]	ANN	SCADA	Gearbox	Monitoring	NA
[24]	Multiple	SCADA	Anomaly	Detection	NA
[30]	LSTM	SCADA	Blade Icing	Detection	NA
[27]	ANN	SCADA	Anomaly	Identification	NA
[12]	Multiple	SCADA	Faults	Prediction	60 min
[13]	Multiple	SCADA	Faults	Monitoring	0
[29]	Kohonen Network	SCADA	Anomaly	Detection / Prediction	1 day

Table 1: Literature on machine learning applied to WTG failure.

Data Sample Structures

Input Features

Operational Data | Meteorological Data

Measurements

Time of Inference

Forecasts

Time UTC	S_1	S_2	…	Sn	Time UTC	M_1	M_2	…	M_k
12:00	$x_{1,1}$	$x_{1,2}$	…	$x_{1,n}$	12:00	$x_{1,n+1}$	$x_{1,n+2}$	…	$x_{1,n+k}$
12:10	$x_{2,1}$	$x_{2,2}$	…	$x_{2,n}$	12:10	$x_{2,n+1}$	$x_{2,n+2}$	…	$x_{2,n+k}$
12:20	$x_{3,1}$	$x_{3,2}$	…	$x_{3,n}$	12:20	$x_{3,n+1}$	$x_{3,n+2}$	…	$x_{3,n+k}$
Time UTC	**S_1**	**S_2**	**…**	**S_n**	**Time UTC**	**M_1**	**M_2**	**…**	**M_k**
11:40					11:40	$x_{145,n+1}$	$x_{145,n+2}$	…	$x_{145,n+k}$
11:50	***Unknown***				11:50	$x_{146,n+1}$	$x_{146,n+2}$	…	$x_{146,n+k}$
12:00					12:00	$x_{147,n+1}$	$x_{147,n+2}$	…	$x_{147,n+k}$

Output Labels

Label:

$$y_t \in \{0, 1\}$$

0 means no event
1 means icing event

t is the time horizon
e.g. $t = 1$ day ahead

Inference:

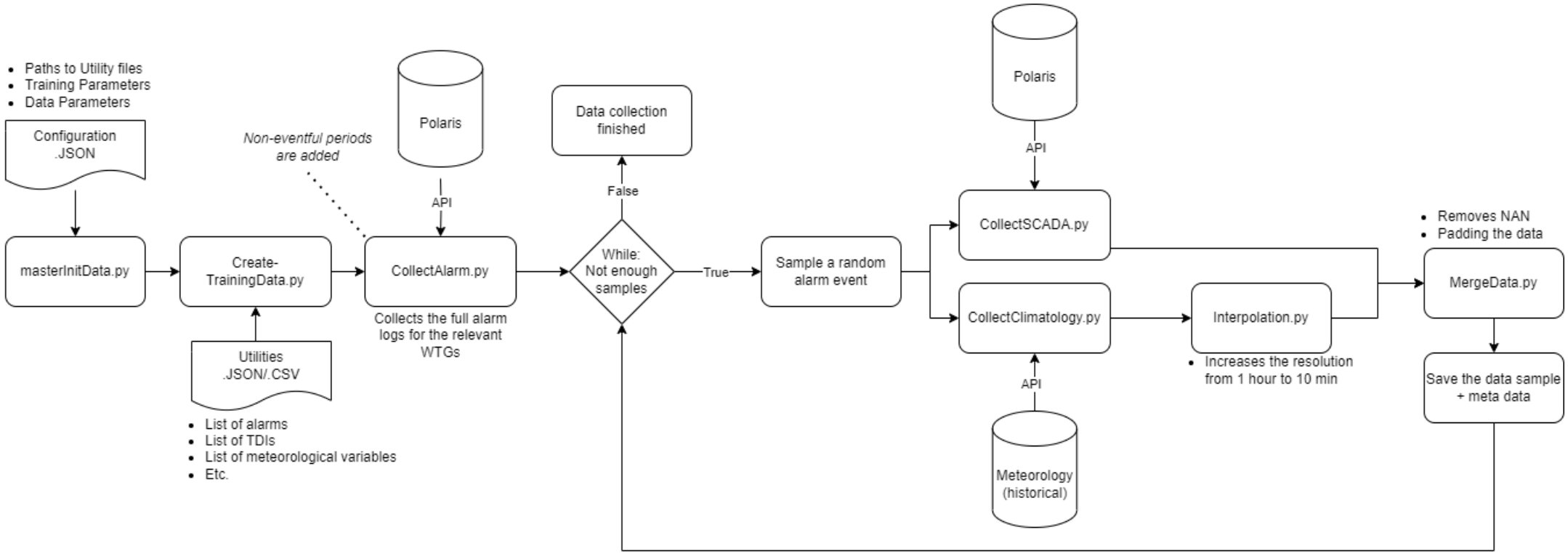
$$\hat{y}_t \in [0, 1]$$

Sampling Process and Training Data Collected

Data used for training:
- ~40.000 samples (769 years)
- ½ are icing events
- ½ without icing events

Data used for validation:
- ~5.000 samples (96 years)
- ½ icing events
- ½ without icing events

Collected from:
- 427 sites
- 1.952 WTGs
- ~ 12 icing events per WTG



Conditional Bilateral LSTM deep learning architecture

Recurrent Neural Network:

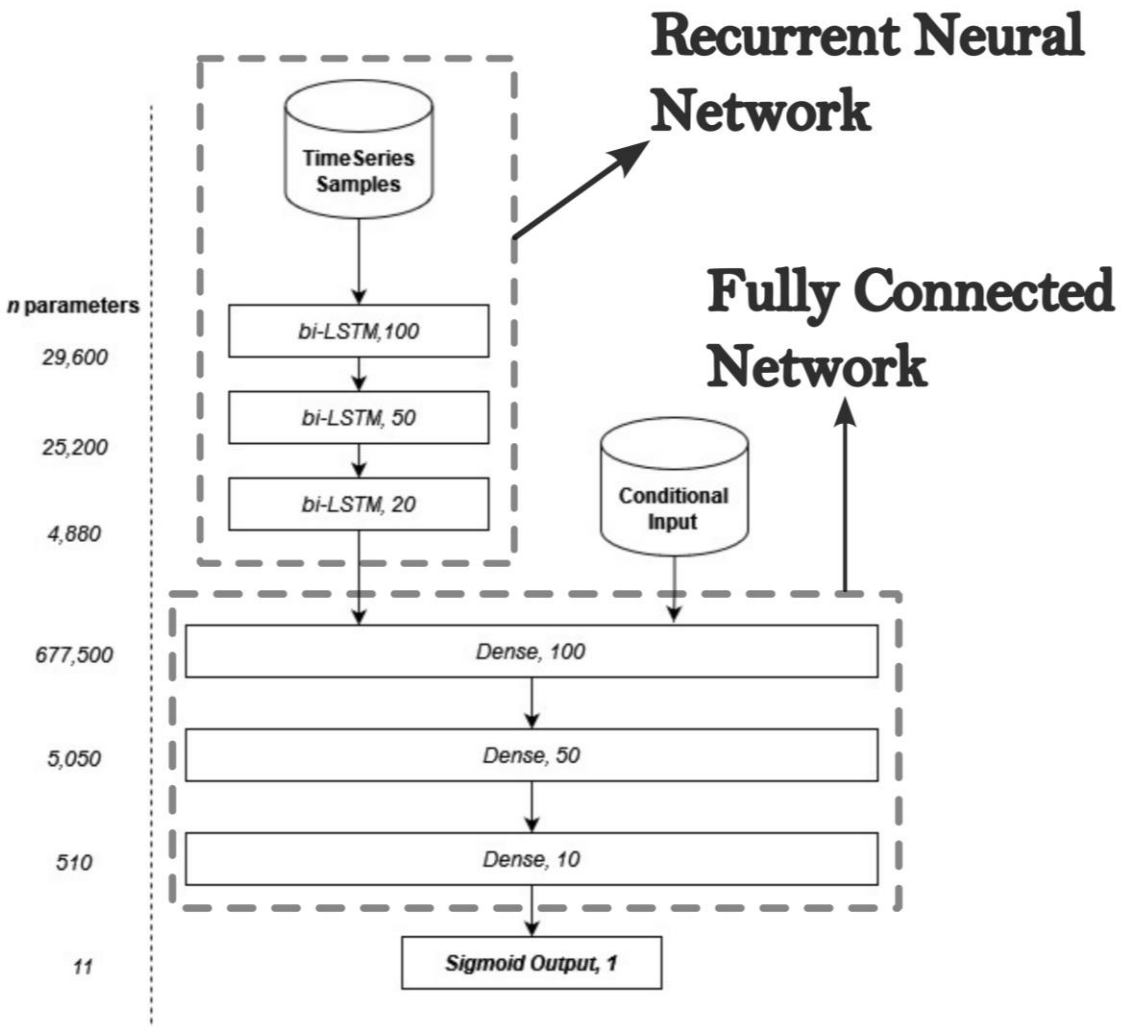
- Feature learning from timeseries
- Parameters: 59.680

Fully Connected Network

- Bottlenecks features from RNN to output
- Uses conditional identifiers to minimize concept drift
- Parameters: 683.071

Output Layer

- Sigmoid for binary predictions
- Threshold of 0.5
- (Softmax for multiple output predictions)



Results and Confusion Matrix

Performance Metrics

F1 Score	0.7386
Accuracy	0.7704
Sensitivity	0.7299
Specificity	0.8028
Precision	0.7476
Negative Predictive Value	0.7878
False Positive Rate	0.1972
False Discovery Rate	0.2524
False Negative Rate	0.2701
Matthews Correlation Coefficient	0.5341

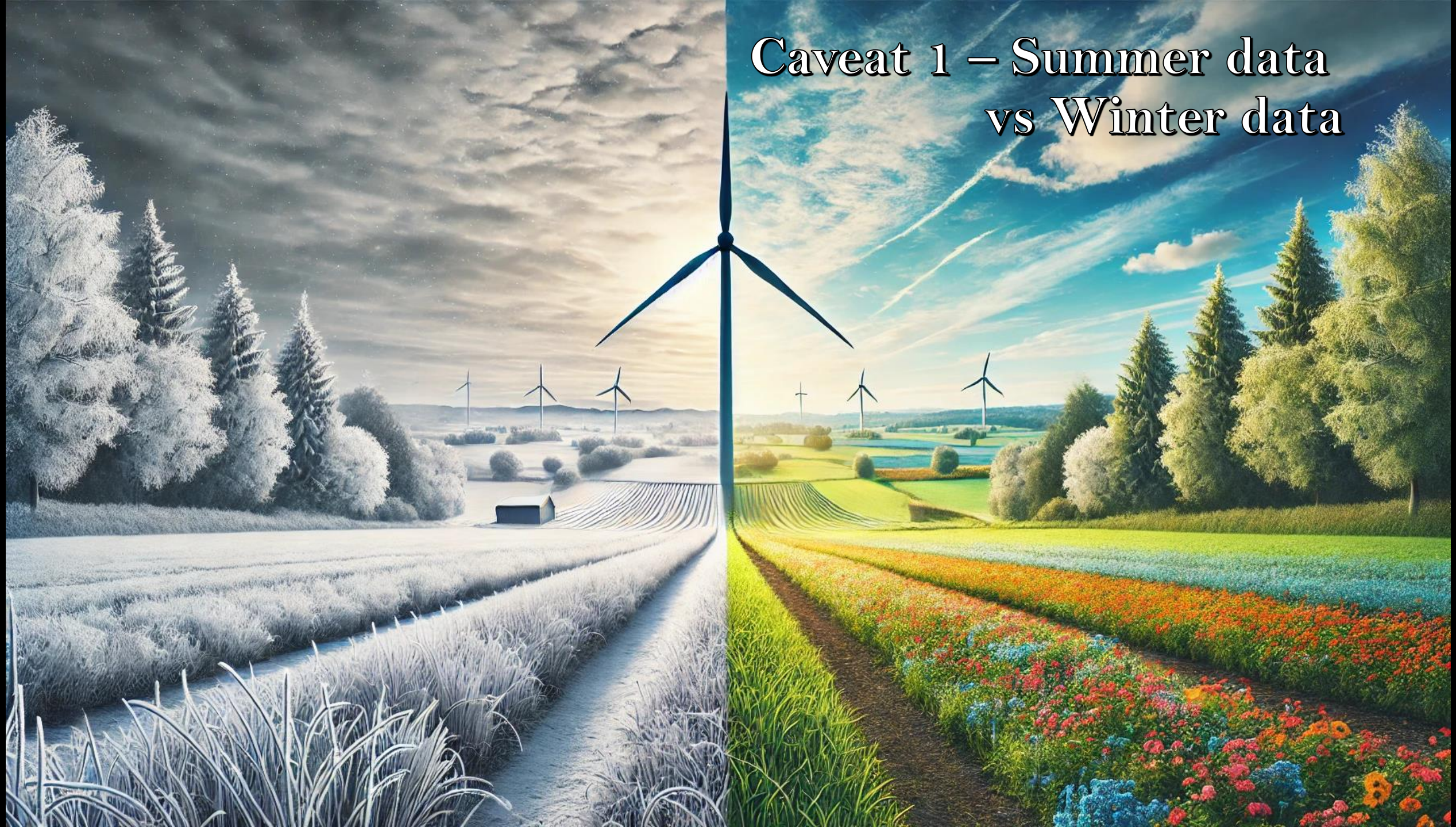
No forecast data used…

Training Process:

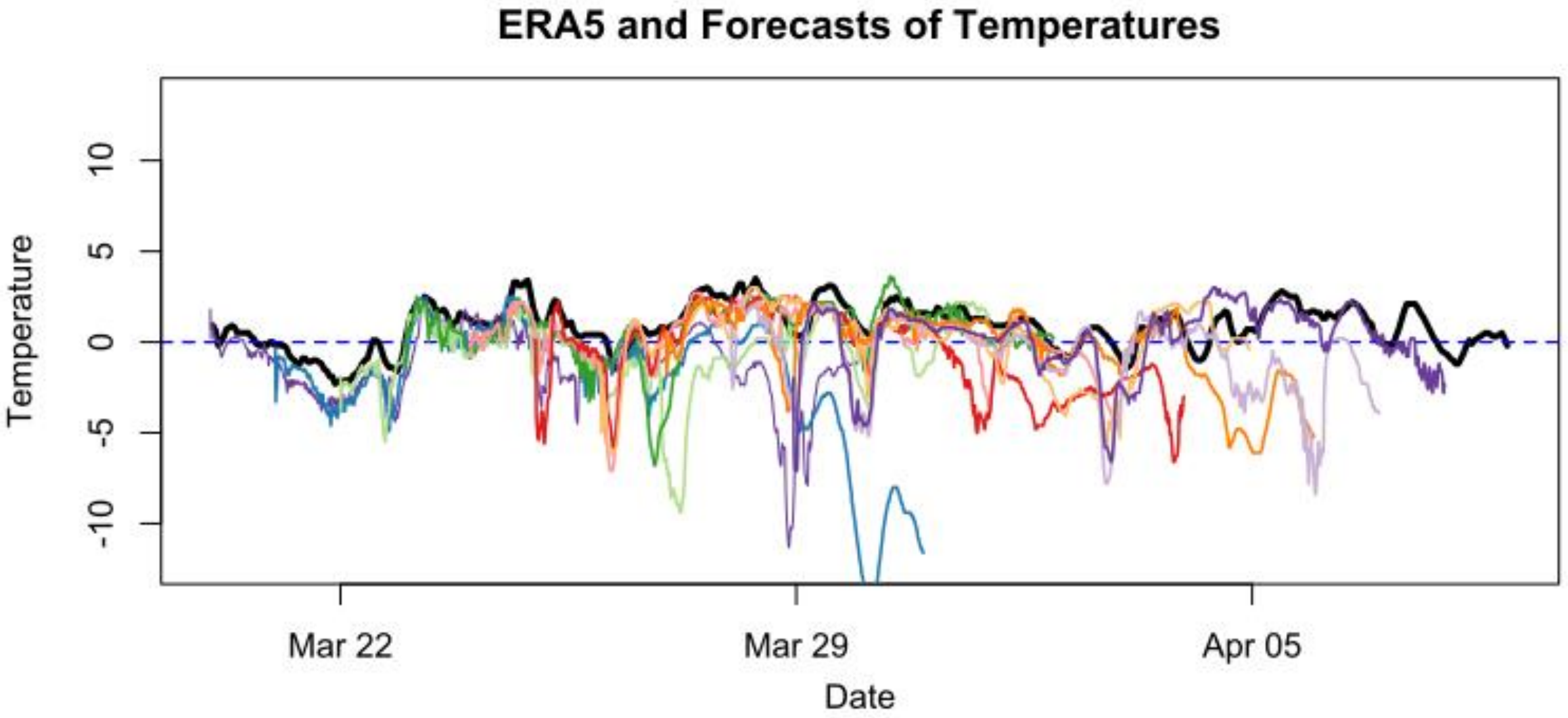
- 70 minutes training time
  - On a *laptop*!
  - 15 epochs
  - Many hyperparameters tested

TARGET / OUTPUT	Class0	Class1	SUM
Class0	2369 44.59%	582 10.95%	2951 80.28% 19.72%
Class1	638 12.01%	1724 32.45%	2362 72.99% 27.01%
SUM	3007 78.78% 21.22%	2306 74.76% 25.24%	4093 / 5313 77.04% 22.96%

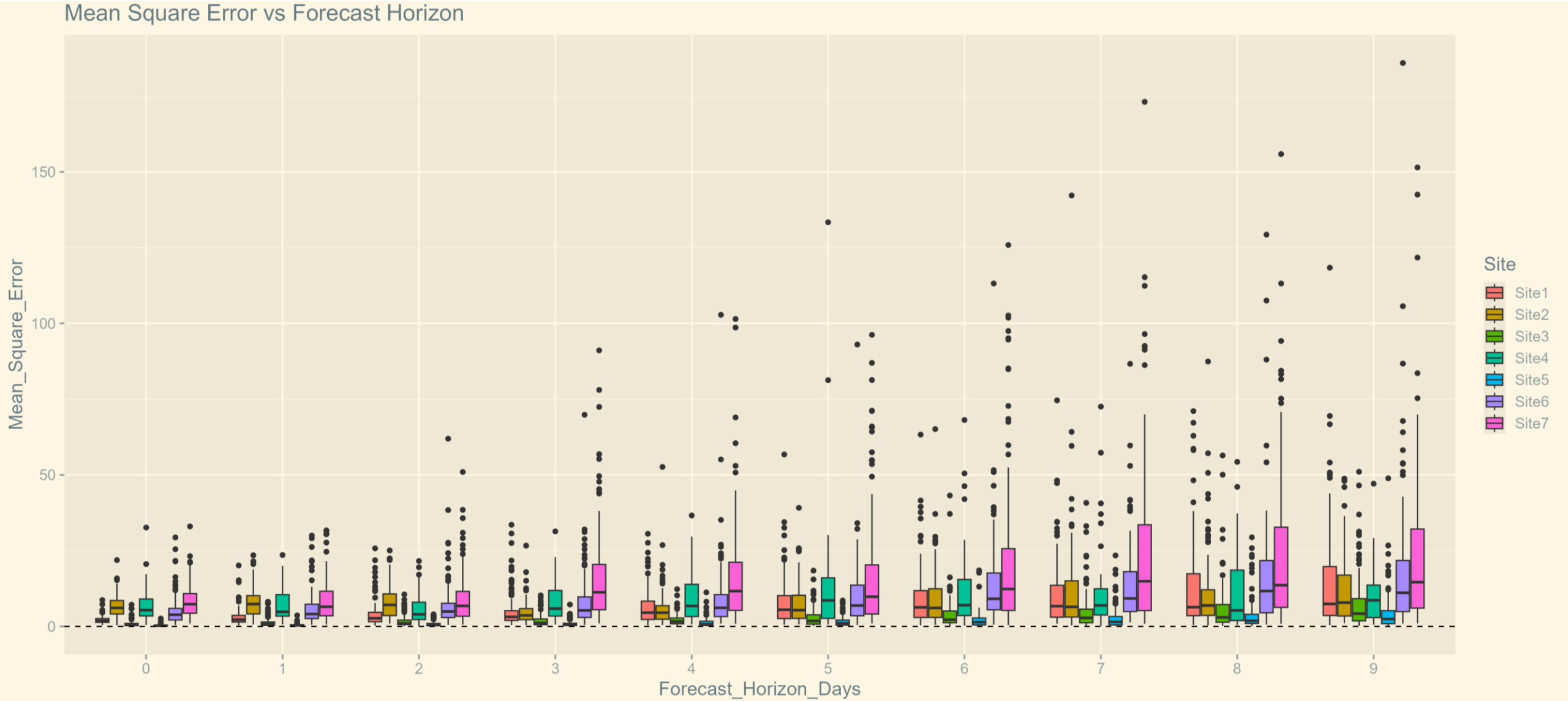
74% of Icing events are predicted


Caveat 1 – Summer data
vs Winter data

Caveat 2 – Weather Forecast Uncertainty



Caveat 2 – Weather Forecast Uncertainty



Caveat 3 – Concept Drift and Dissimilar WTGs

In the Training data

Variable	# items
Countries	31
Sites	427
WTG Types	33
WTG Platforms	5
Nominal Power	1.3 – 8.0

Out of 1952 WTGs…

Higher *dissimilarity* in training data leading to degrading model performance

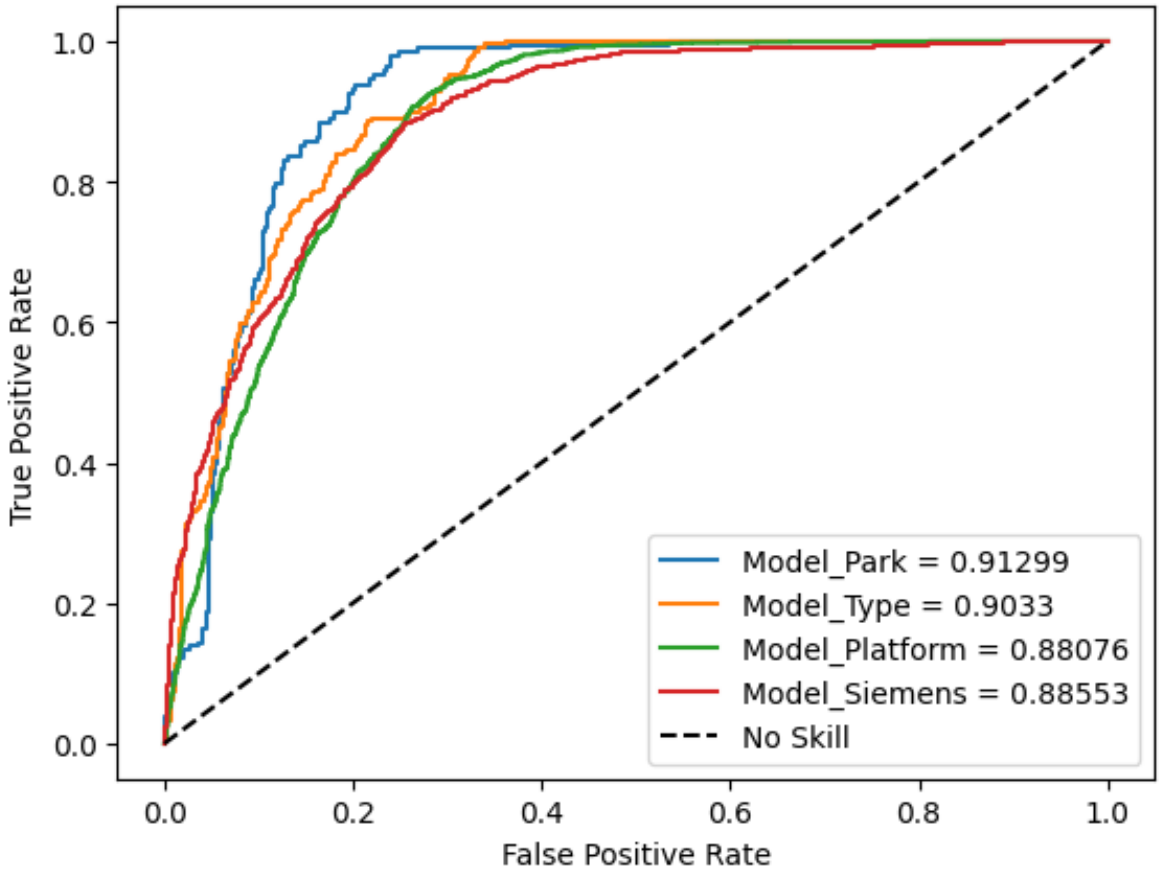
Increase Data Size

Decrease Dissimiliarity

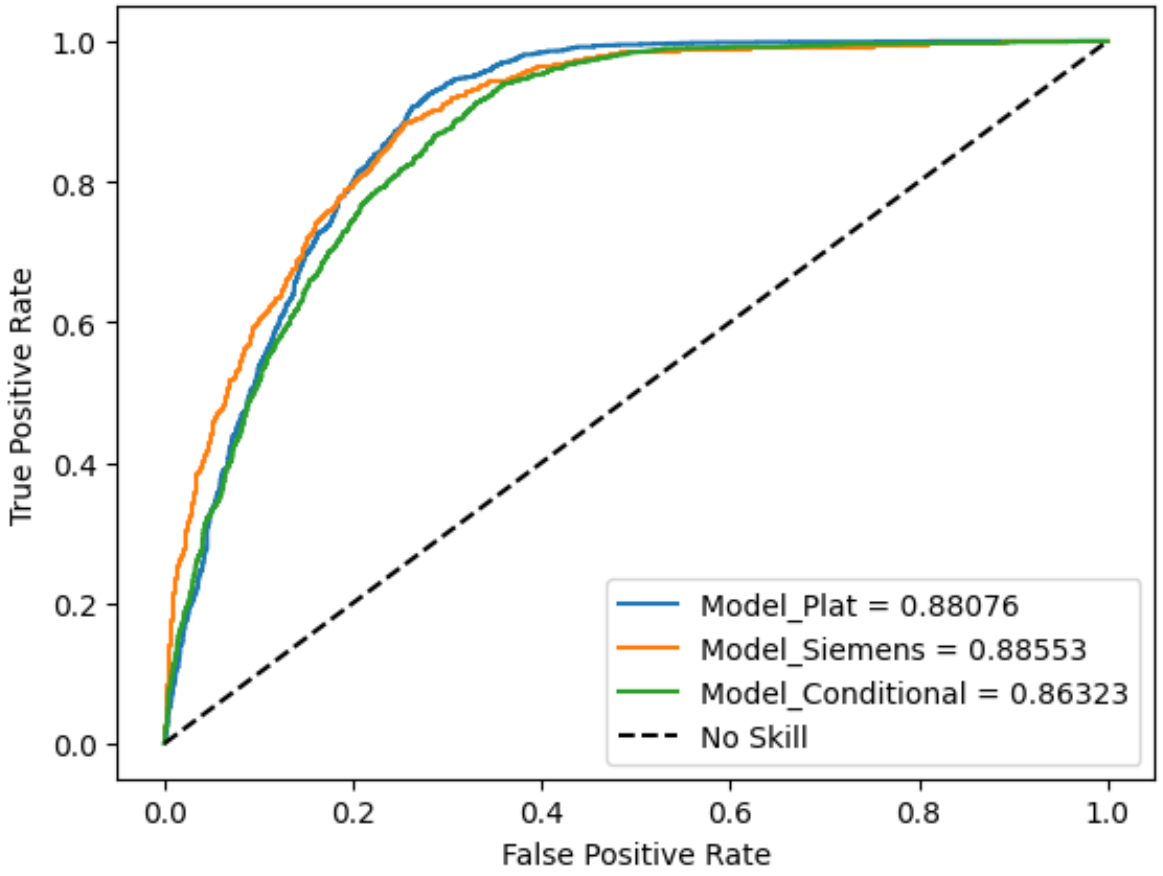
Less training data and less applicability in deployment

Caveat 3 – Concept Drift and Dissimilar WTGs

Non-Conditional Model



Conditional Model



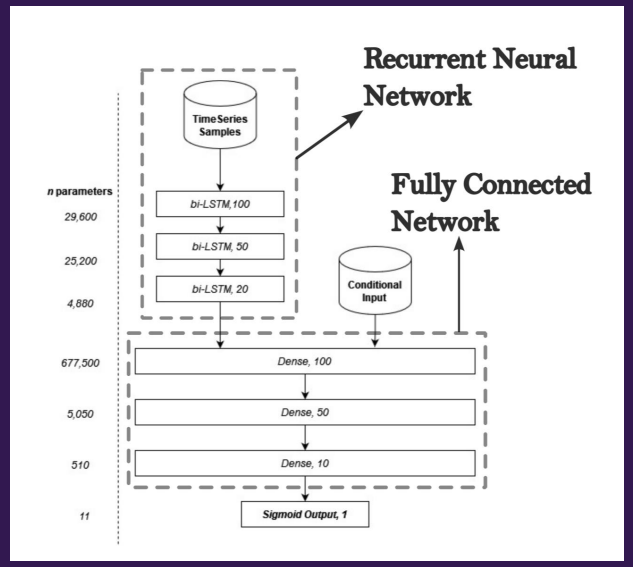
Summary and conclusions

1. Problem

- Loss of operational hours because of icing events

2. Method

- Data sources and samples
- Deep learning architecture



3. Results

- F1 - 0.74
- Accuracy – 0.77

TARGET / OUTPUT	Class0	Class1	SUM
Class0	2369 44.59%	582 10.95%	2951 80.28% 19.72%
Class1	638 12.01%	1724 32.45%	2362 72.99% 27.01%
SUM	3007 78.78% 21.22%	2306 74.76% 25.24%	4093 / 5313 77.04% 22.96%

74% of Icing events are predicted

4. Future work

Increase Data Size

Decrease Dissimiliarity