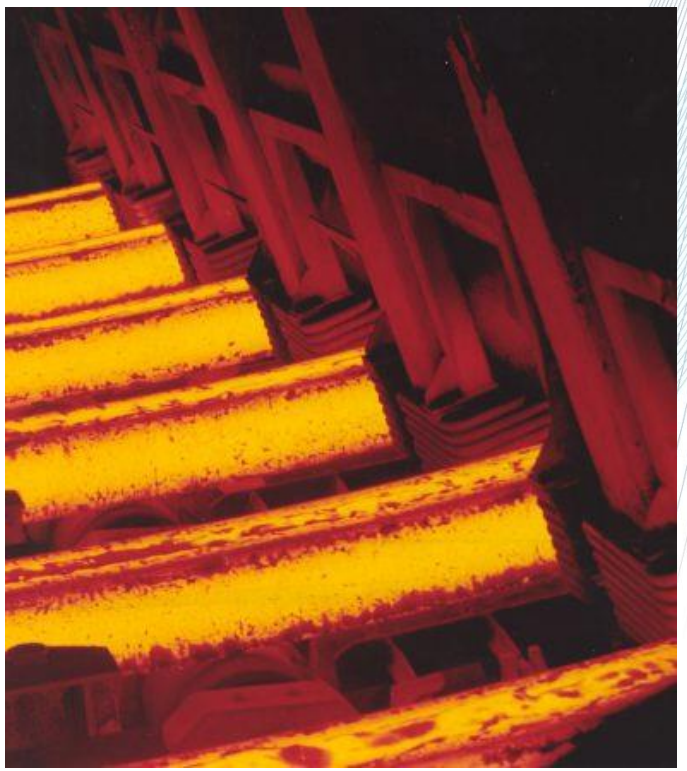
# RealTimeCastSupport

## Embedded real-time analysis of continuous casting for machine-supported quality optimisation

Webinar on 8th of September 2023

**Exploitation of various CC data**

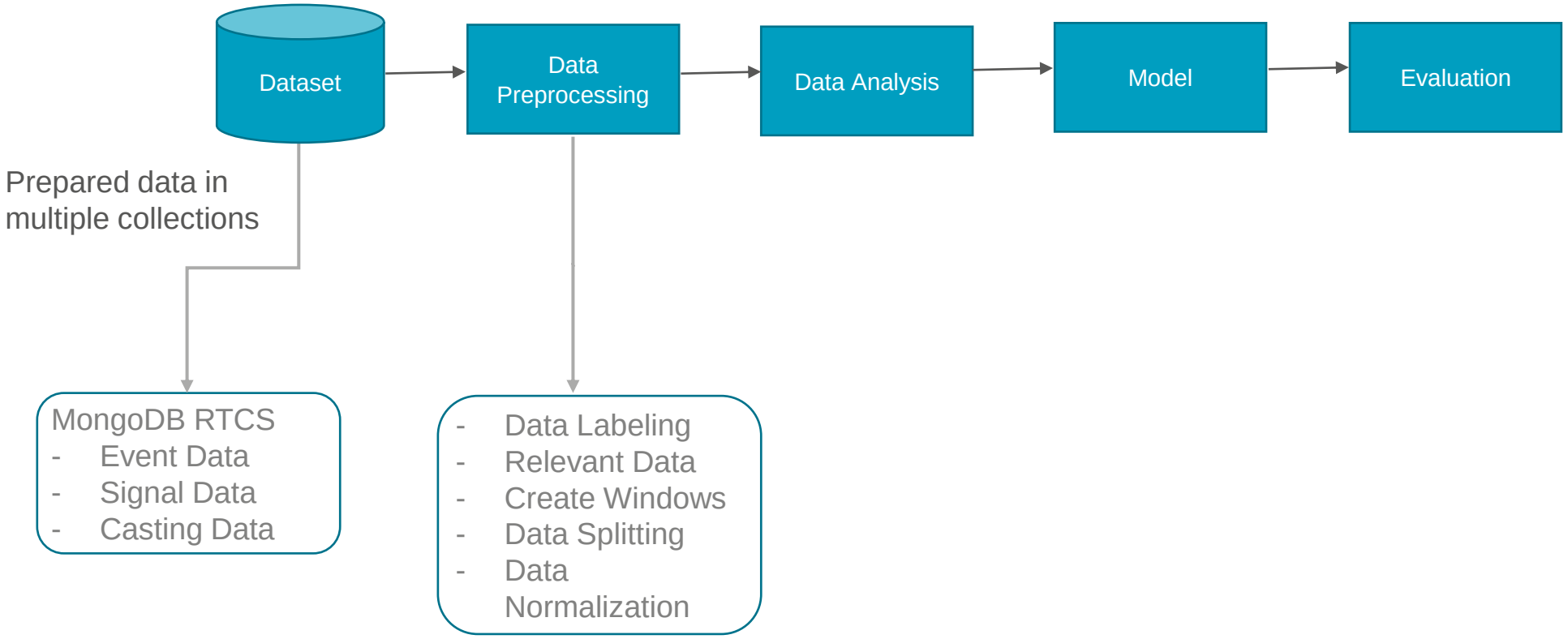
Zeinab Kargar (BFI)



BFI Excellence in Applied Research | VDEh-Betriebsforschungsinstitut GmbH

Stahl

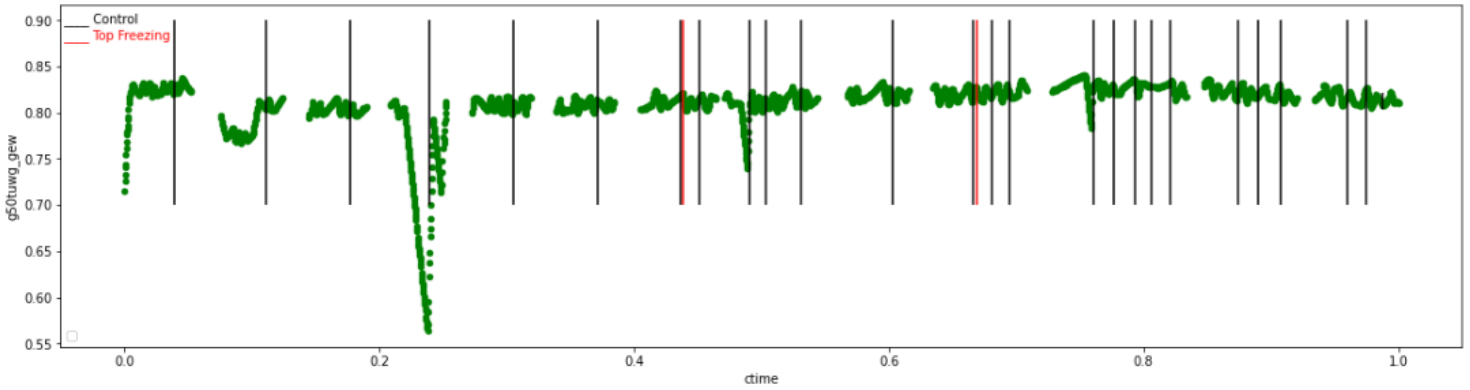
# Framework for Data Analysis



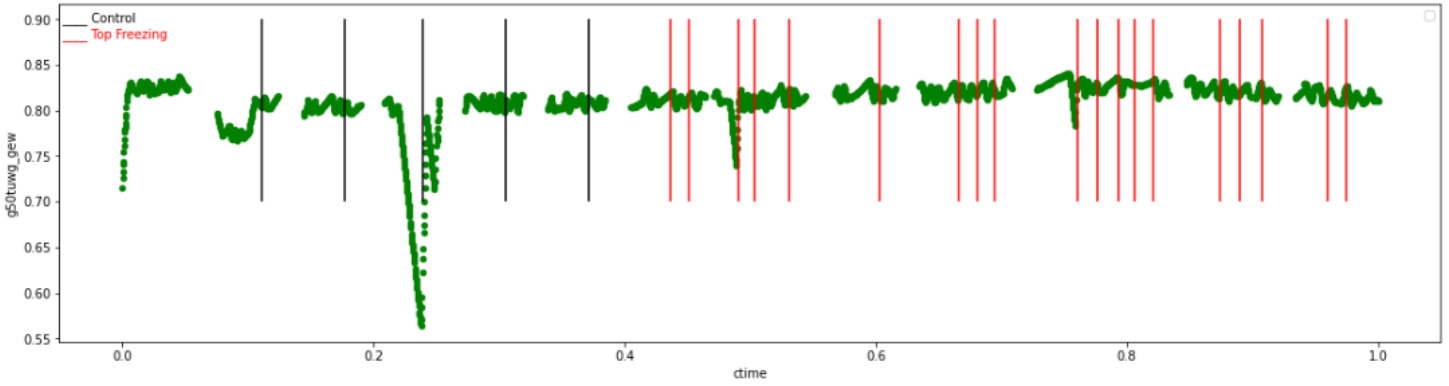
# Data Preprocessing

Data Labeling

| Code | Number of events<br>Plant CC5 | Number of events<br>Plant CC4 |
|---|---|---|
| Control strand 1 | 7958 | 4968 |
| Start top freezing strand 1 | 8 | 12 |
| End top freezing strand 1 | 1 | 8 |
| Control strand 2 | 4968 | 842 |
| Start top freezing strand 2 | 8 | 11 |
| End top freezing strand 2 | 5 | 11 |




| Code | Number of events<br>Plant CC5 | Number of events<br>Plant CC4 |
|---|---|---|
| Control strand 1 | 7929 | 4904 |
| Start top freezing strand 1 | 131 | 77 |
| Control strand 2 | 2015 | 827 |
| Start top freezing strand 2 | 33 | 34 |



# Data Preprocessing

Relevant Time Signals

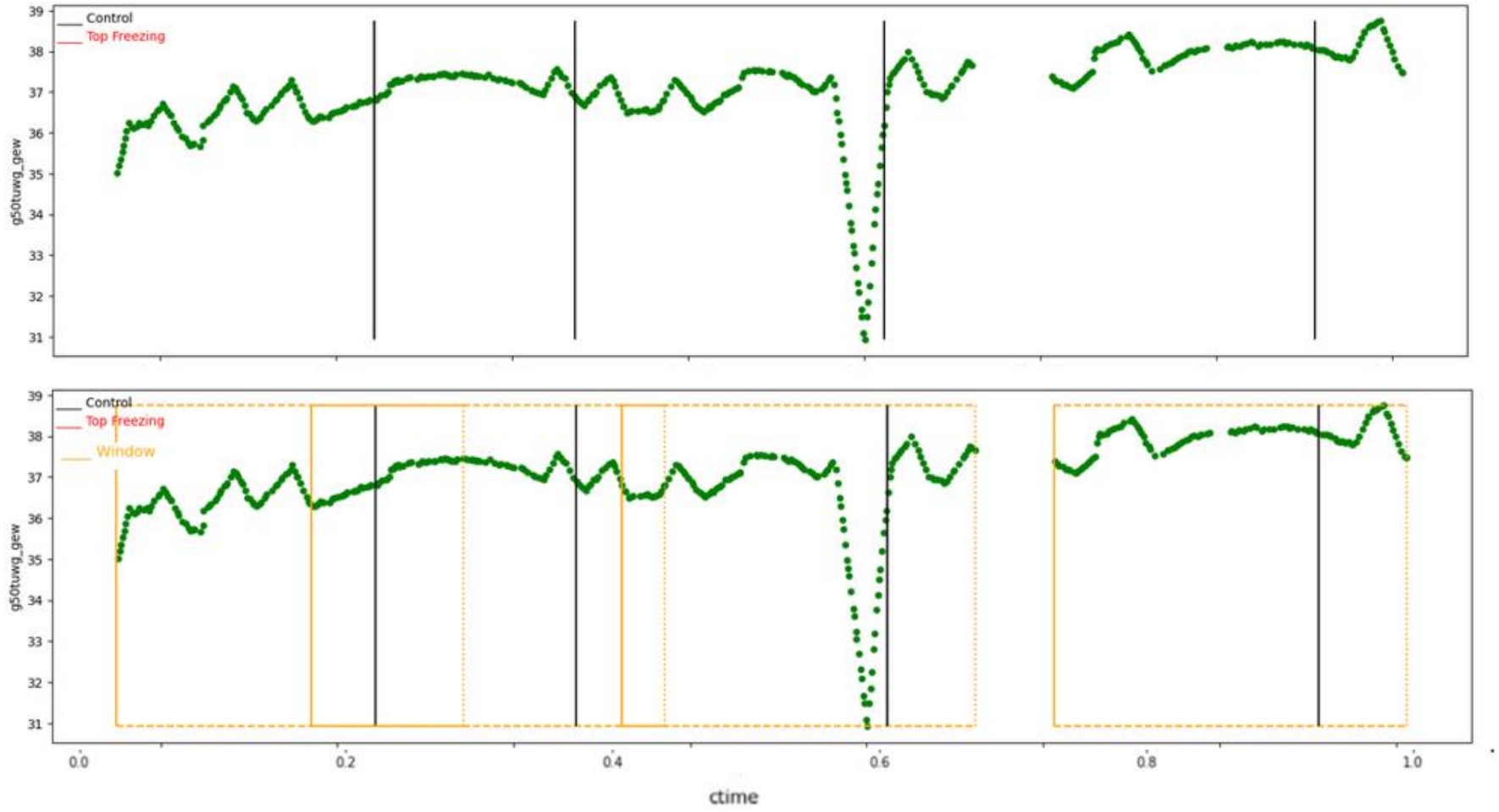
**Criteria for selecting signals**

-Continuous casting experts
-CFD Model results

| Sensor name | Description | Unit | 1s or worse |
|---|---|---|---|
| g50tu_fuell | Tundish filling level | mm | x |
| g50tuwg_gew | Tundish weight | t | x |
| g50tu_cycle | Tundish cycle (Distance Tundish-Mould) | mm | x |
| g51tusp_sgf | TundishPlug_Purge GasFlow | l/min | x |
| 'g51tusp_stell' | Stopper position | mm | x |
| g51_giv | Casting speed | m/min | x |
| g51_seql | Sequence length | m | x |
| g51ko_fuell | Mould level | mm | x |
| g51kodf_wsd | Heat flux density Dieflen | MW/m2 (calculated) | x |
| g51kodl_wsd | Heat flux density Dillingen | MW/m2 (calculated) | x |
| g51kofs_wsd | Heat flux density fixed side | MW/m2 (calculated) | x |
| g51kols_wsd | Heat flux density loose side | MW/m2 (calculated) | x |
| g51ko_hubfrq | Oscillation frequency | oscillations/min | x |
| g51ko_gipuvb | Casting powder consumption | kg/t | x |

# Data Preprocessing

Create Window



# Data Preprocessing

Data Splitting

80% for the training data 20% for the test data
all top freezing events are taken as test set

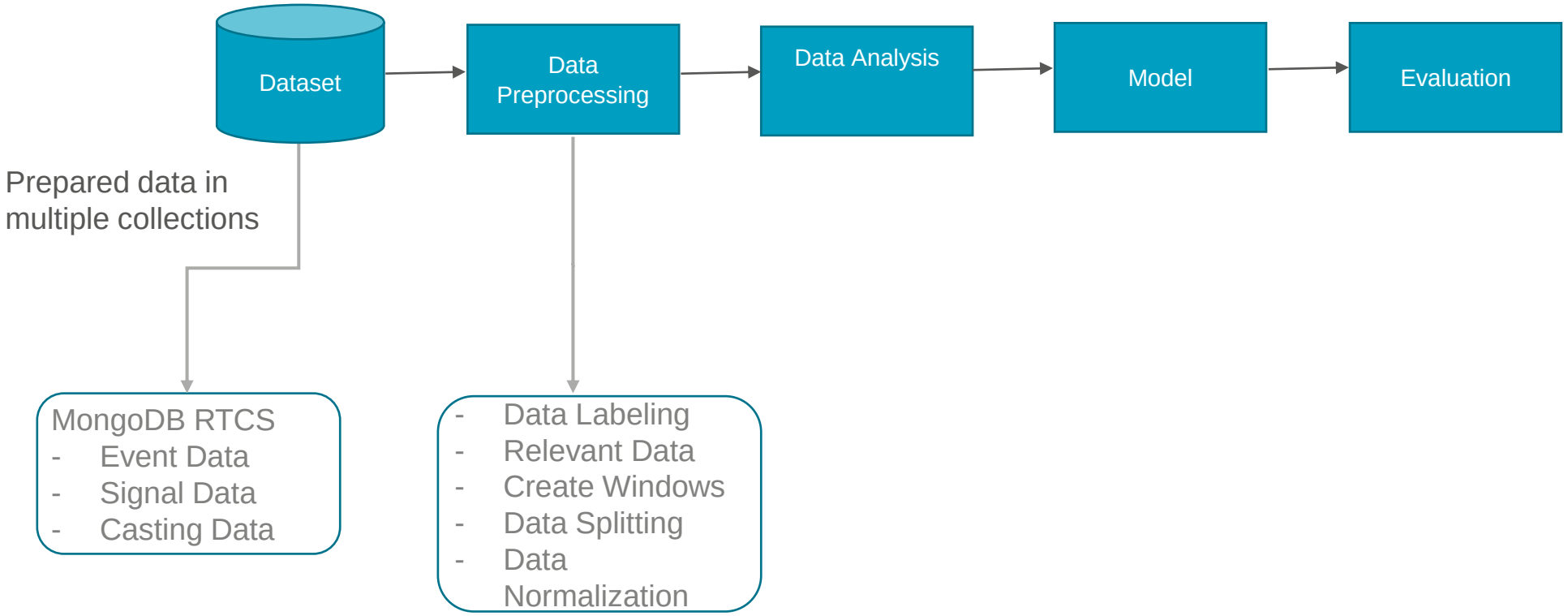
Data Normalization

- The signals of continuous casting have different ranges
- Rescale input signals prior to training machine learning model

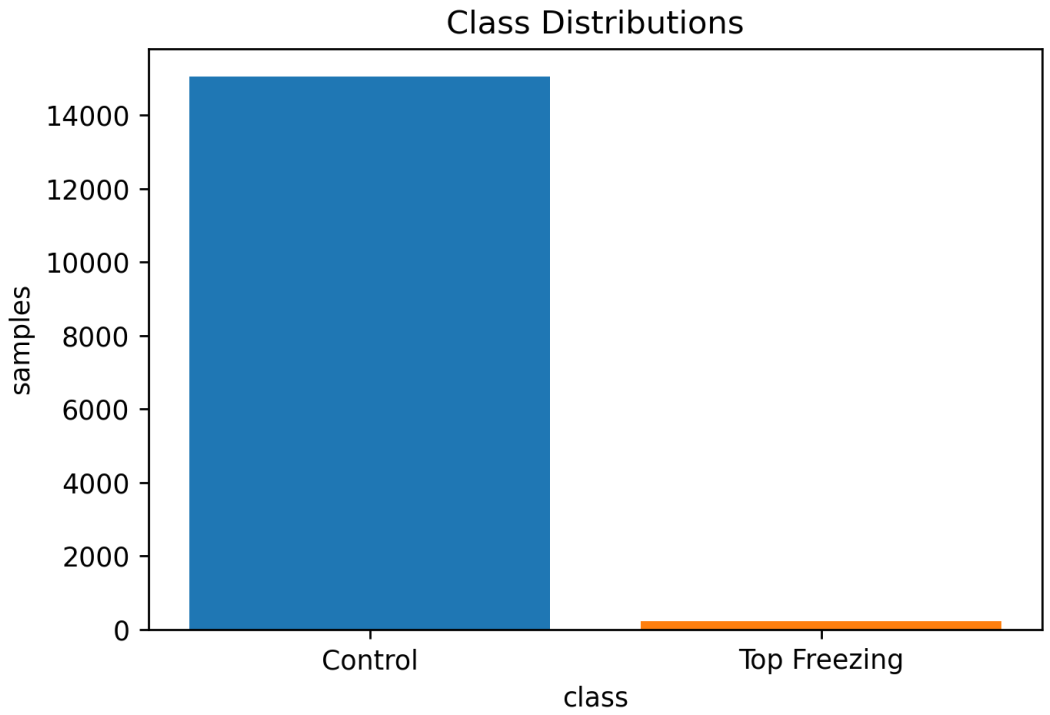
Min-Max Scaler

$$X' = \frac{X - X\_min}{X\_max - X\_min}$$

# Framework for Data Analysis



# Data Analysis



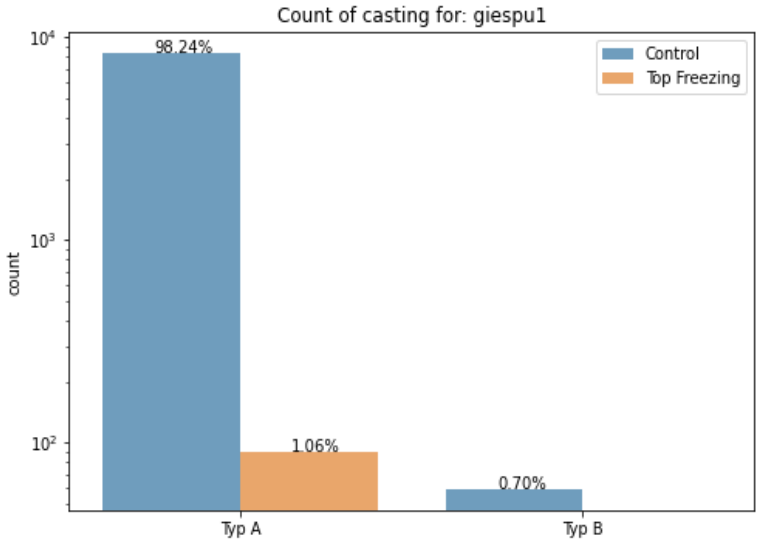

Highly Imbalanced Dataset

The percentage of samples:

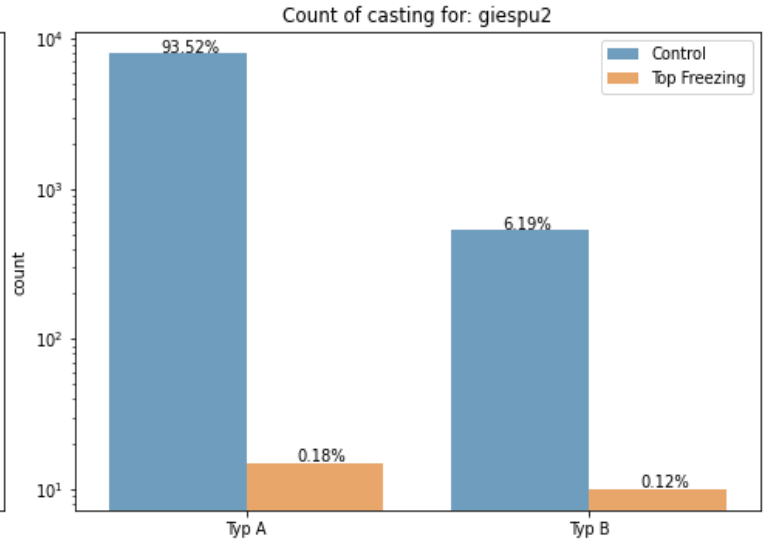
Control class: 98.55 %
Top freezing class: 1.44 %

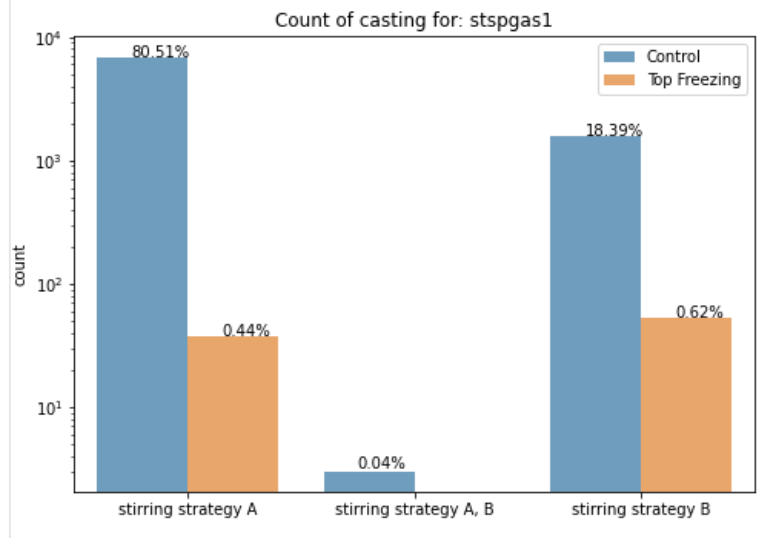
# Data Analysis

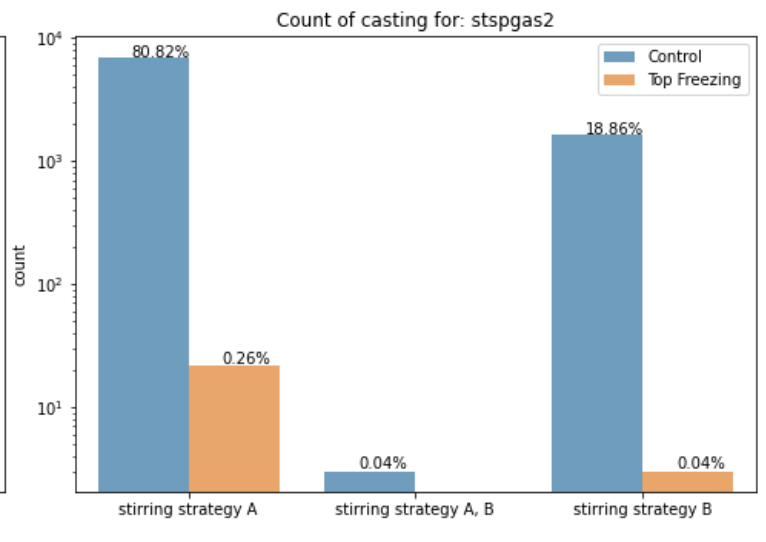
Strand 1




Strand 2



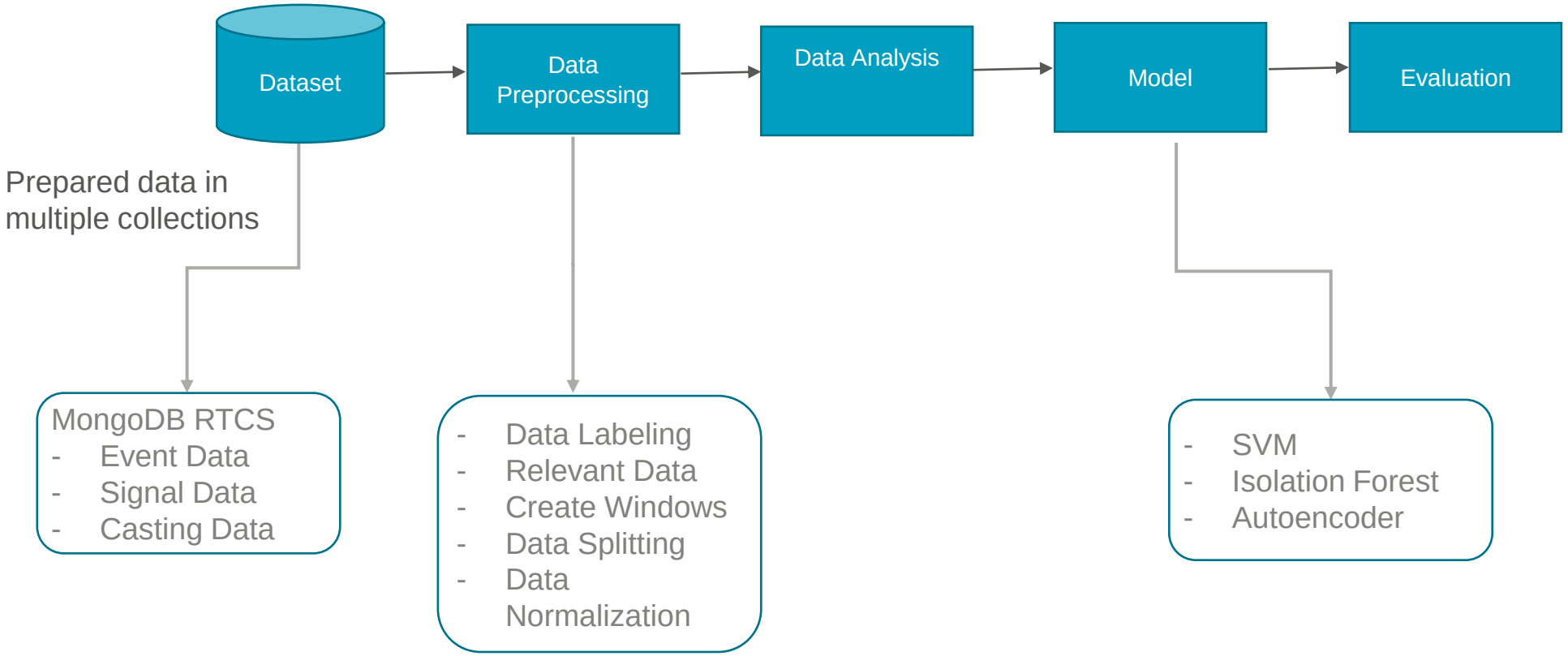




## Casting Powder

Type A has a lower carbon content than Type B

## Gas Injection

# Framework for Data Analysis



# Modeling

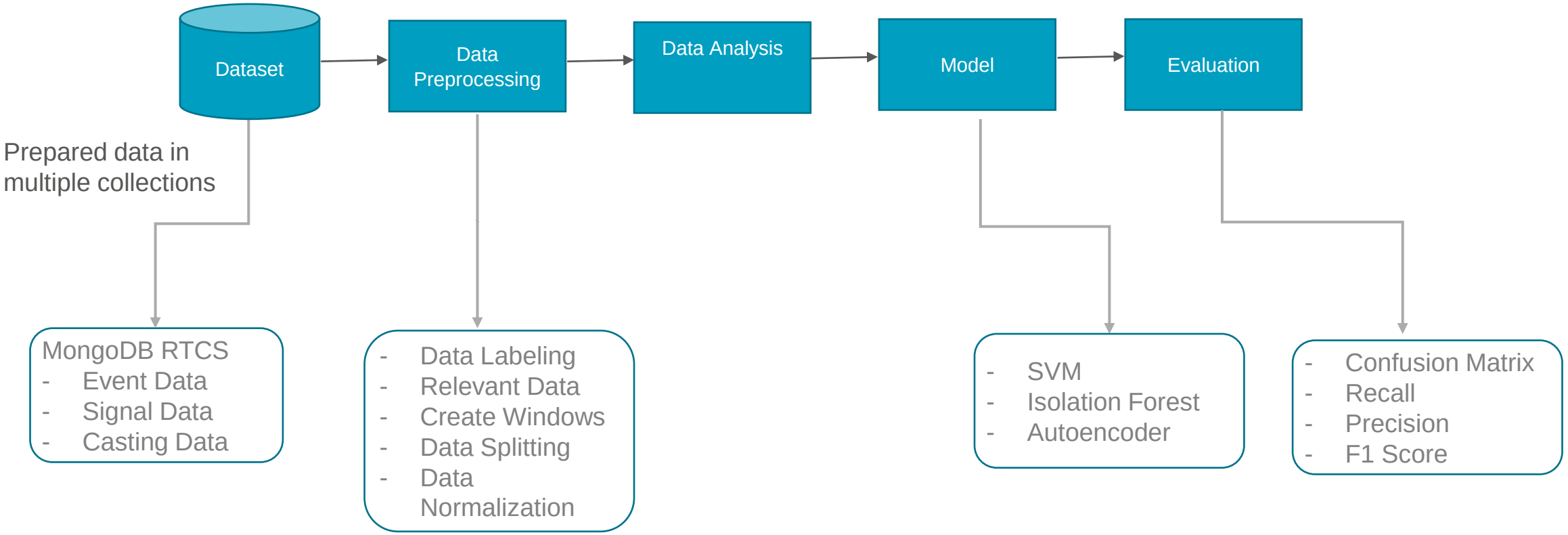
- Identification of rare events (Anomaly detection )

- Highly imbalanced dataset

- One-class classification Machine learning models Use a normal training dataset to create a model that represents normal behavior. Anomalies can then be detected by deviating from this model

- One-Class Support Vector Machines - (SVM)
- Isolation Forest
- Autoencoder

# Framework for Data Analysis



# Model evaluation

Confusion Matrix

| Acctual Values | | Predicted Values: Negative | Predicted Values: Positive |
|---|---|---|---|
| | Negative | TN<br>True Negative | FP<br>False Positive |
| | Positive | FN<br>False Negative | TP<br>True Positive |

**Accuracy** : The number of correctly classified data over the total number of data.

$$Accuracy = \frac{TN + TP}{TN + FP + FN + TP}$$

**Recall**: What proportion of true anomalies was identified ?

$$Recall = \frac{TP}{TP + FN}$$

**Precision** What proportion of identified anomalies are true anomalies?

$$Precision = \frac{TP}{TP + FP}$$

**F1 Score** : Combining both Recall and Precision

$$F1\ Score = \frac{2 * Recall * precision}{Precision + Recall}$$

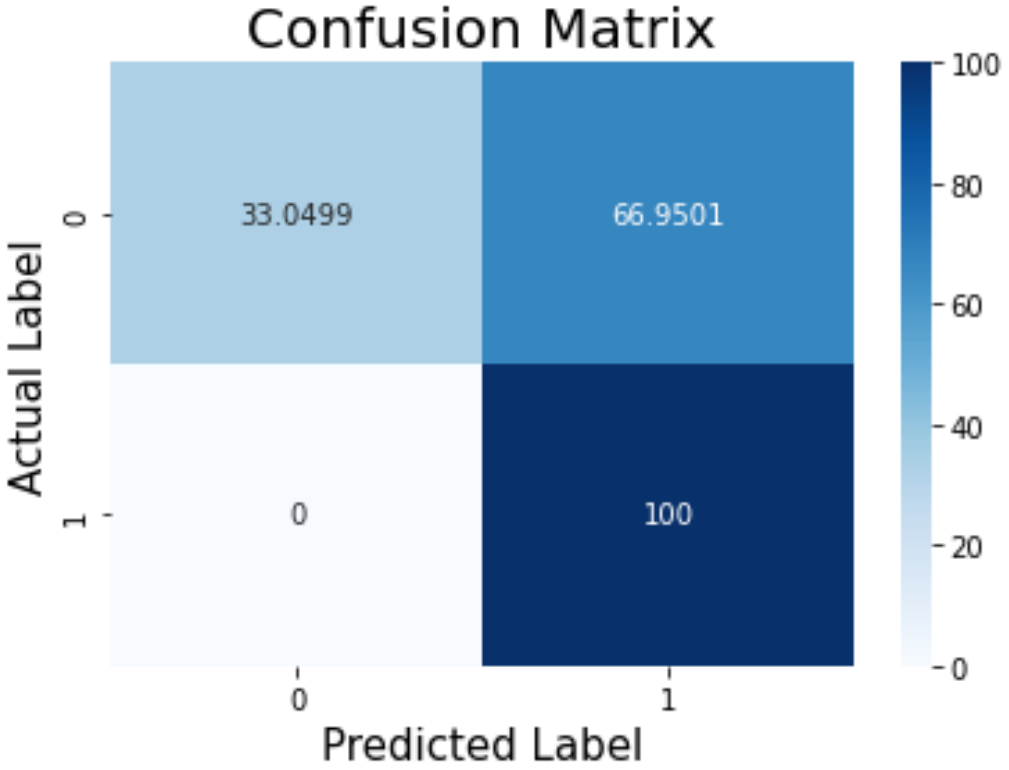
# Results

| Model Name | Data set | Accuracy | Recall | Precision | F1_score |
|---|---|---|---|---|---|
| SVM | CC5/strand1 | 0.70 | 0.48 | 0.06 | 0.10 |
| ISF | CC5/strand1 | 0.69 | 0.25 | 0.03 | 0.05 |
| AE | CC5/strand1 | 0.92 | 0.02 | 0.02 | 0.02 |
| | | | | | |
| SVM | CC5/strand2 | 0.70 | 0.58 | 0.04 | 0.089 |
| ISF | CC5/strand2 | 0.69 | 0.25 | 0.02 | 0.03 |
| AE | CC5/strand2 | 0.81 | 0.13 | 0.25 | 0.17 |
| | | | | | |
| SVM | CC4/strand1 | 0.69 | 0.25 | 0.033 | 0.05 |
| ISF | CC4/strand1 | 0.68 | 0.27 | 0.03 | 0.062 |
| AE | CC4/strand1 | 0.90 | 0.20 | 0.27 | 0.23 |
| | | | | | |
| SVM | CC4/strand2 | 0.62 | 0.5 | 0.19 | 0.28 |
| ISF | CC4/strand2 | 0.61 | 0.46 | 0.18 | 0.26 |
| AE | CC4/strand2 | 0.81 | 0.16 | 0.29 | 0.22 |

# Results

Hyperparameter Tuning: controlling the behavior of a machine learning model.

One-class SVM, all top freezing events is detected by choosing nu=0.7 for CC5 and strand 1.

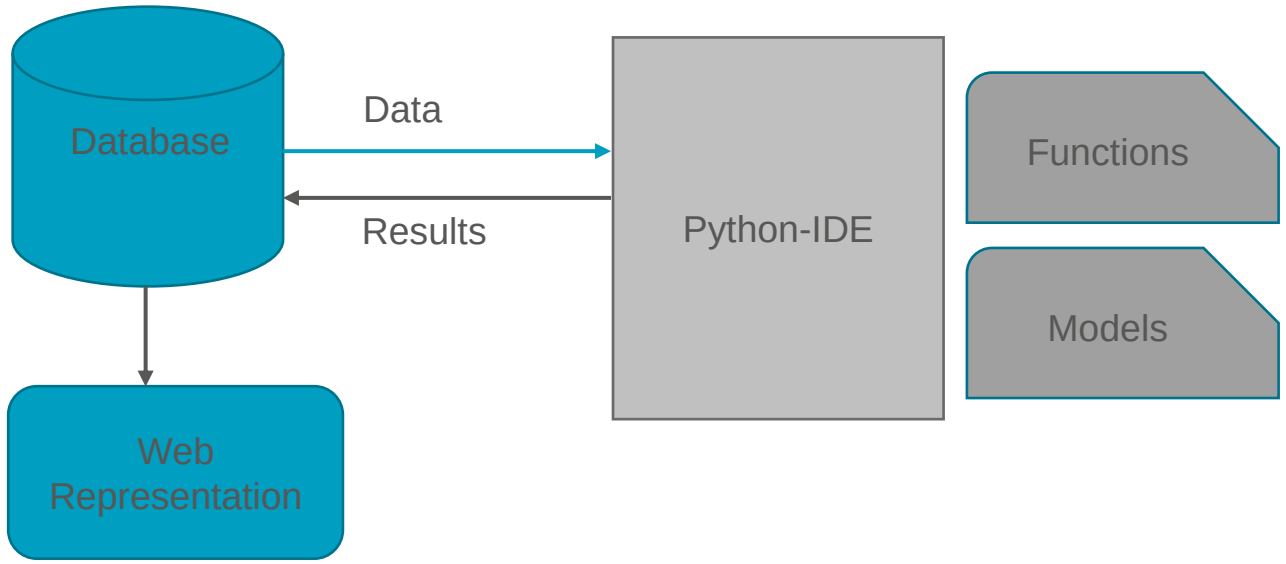


# Next step

Deployment of data-driven methods in real-time

**Potential challenges in real-time implementation:**

- Data integration and pre-processing
- Data quality
- Model drift



# Conclusion

- Leveraged critical casting parameters to gain data-driven insights into top freezing events during continuous casting.
- Challenge with imbalanced data distribution
- Explored a range of machine learning models, including SVM, Isolation Forest, and Autoencoder
- Evaluated model effectiveness using the 'recall' metric.
- By optimizing the machine learning models, the detection of true positives was increased at the expense of false positives.