SEMANTIC AI FOR PREDICTIVE MAINTENANCE OF RAILWAY TRACK SYSTEMS

Edmond S. H. CHEONG (lead presenter)[1], Kinson C. H. LAM[2], and Leo K. S. LEE[2], Winson F.S. TSE[3]
Yu YANG[4], S. Joe QIN[4], Qingpeng ZHANG[4], Lishuai LI[4] and Paul H. F. LAM[5]

[1] *Electrical and Mechanical Services Department, the Government of the Hong Kong Special Administrative Region, 3 Kai Shing Street, Kowloon, Hong Kong*
(Email: shcheong@emsd.gov.hk, +852-3912 0623)

[2] *Electrical and Mechanical Services Department, the Government of the Hong Kong Special Administrative Region, 3 Kai Shing Street, Kowloon, Hong Kong*
(Email: kinson@emsd.gov.hk, +852-3912 0612 / leekaishing@emsd.gov.hk, +852-3912 0620)

[3] *General Manager – Infrastructure Maintenance, MTR Corporation, Fo Tan Railway House, Fo Tan, N.T., Hong Kong.*
(Email: twinson@mtr.com.hk, +852-2688 1337)

[4] *Hong Kong Institute for Data Science / School of Data Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong*
(Email: yuyang@cityu.edu.hk, Joe.Qin@cityu.edu.hk, qingpeng.zhang@cityu.edu.hk, lishuai.li@cityu.edu.hk)

[5] *Department of Architecture and Civil Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong*
(Email: paullam@cityu.edu.hk)

BACKGROUND

Hong Kong is one of the most densely populated cities in the world. To meet the citizens' travel needs, the railway operates with long daily service hours, resulting in a short maintenance window. To further enhance the safety and reliability of railway, it is important to implement effective approaches to analyze railway incidents, explore correlations of incidents, and recommend alerts of high-risk equipment and areas in railway systems to achieve novel predictive maintenance on railway track systems.

Recent years have witnessed a sharp penetration of data-driven methods into various industry sectors, such as manufacturing, finance, transportation, cybersecurity, and healthcare. Therefore, in this study, we aim to exploit a wide range of railway data, such as the railway incident reports, maintenance records, real-time condition data and online information, which are valuable for gaining insights into associated factors with different degrees of connection in leading to railway track incidents, by building an Artificial Intelligence (AI) model.

In terms of the AI model building, it is crucial to turn the raw data into structured knowledge so that AI technologies can be used to efficiently process the data and make predictions using the data. Many information extraction techniques are used to transform raw data into structured data. For example, Regular Expressions (Regex) are often used to extract specific keywords; interpolation methods can be applied to handle missing values in the raw data; Interquartile Range (IQR) is designed for removing outliers in the raw data that may mislead downstream data analytics tasks; complex unsupervised learning methods such as Principal Component Analysis (PCA) (Pearson, 1901) and K-means (Lloyd, 1982; MacQueen, 1967) are used to extract informative features.

After information extraction from raw data, how to store and retrieve massive amounts of data efficiently is a critical issue. Data modelled in the forms of tabular relations is stored in relational databases like MySQL, Oracle, etc., or in NoSQL databases like MongoDB, Cassandra, etc., if data is modelled in non-relational forms. On the other hand, graph databases, such as Neo4j, have recently become popular since the graph structure carries more semantics and enables more powerful analytics using AI techniques than traditional tabular structures (Chaudhri et al., 2022). Graph databases represent structured knowledge extracted from raw data where nodes represent objects and edges indicate the relationships between objects. By storing the extracted information in a graph database, we can run efficient queries to generate AI model outputs which give extra information on potential insights about area of focus or prescriptive maintenance actions.

In this study, historical data from MTR Corporation (MTR) was used. It was a good start while it was noted that prediction accuracy would need to be further optimized as there are rare past track incident data due to the effective maintenance regime of MTR over the years. The implication that we derived from the current dataset probably requires more incident data to calibrate the significance and relative weightings in the prediction. More incident data may be obtained, e.g. further historical data, to the AI model for learning.

In the long run, once we have an effective representation of the railway data, we can easily apply AI techniques on the data for important downstream applications. For example, to facilitate efficient and effective summarization on all the railway incidents, we can build a question & answering (Q&A) system within the AI model that uses natural language to issue queries over the database. By exploiting Natural Language Processing (NLP) techniques such as Word2Vec (Mikolov, Sutskever, et al., 2013), TextCNN (Zhang & Wallace, 2015), and Fasttext (Joulin et al., 2016), the semantic properties of textual data can be captured and then used in the Q&A process within the AI model. NLP models pre-trained on large corpus of textual data such as Transformer (Vaswani et al., 2017), ELMO (Peters et al., 1802), GPT (Radford et al., 2018), BERT (Devlin et al., 2018) and XLNET (Yang et al., 2019) can further boost the accuracy of the Q&A system. In addition, with the effective representation of the railway data, we can also apply deep neural networks to construct a classification model for railway incident diagnosis.

OBJECTIVE

The objectives of this project are to build a semantic AI model of railway incident data which enables powerful data-driven analytics for discovering knowledge of railway track incidents, interactive Q&A queries, and predictive models of track incidents. We have the following specific objectives: -

1. Build a semantic AI model with “Railway Schema”, “Semantic Model” and “Railway AI Predictive Maintenance Model” that aims to identify risks and predict potential incident occurrences in the railway track system.

2. Integrate a wide range of data sources, including Incident Reports, Track Incident Reports (TIR), Track Geometry Data, Welding Records (WR), Real-time Vibration Data, Abnormal Signal (AS) Records, and Defect Records (DR) of Permanent way (Pway) system, and weather information & news in the public domain, into the

information system. We will turn the diverse data sources into structured knowledge about railway incidents and store the structured knowledge in a clustered database.

3. Build an analytics engine based on the clustered database. By using the analytics engine, we will develop semantic models to identify correlations and contributing factors of railway track incidents based on the Pway data of the operating railway. It would give further insights on area of focus or prescriptive maintenance actions in daily maintenance.

METHODS

In this section, we summarize the methodology used in each main component of the system.

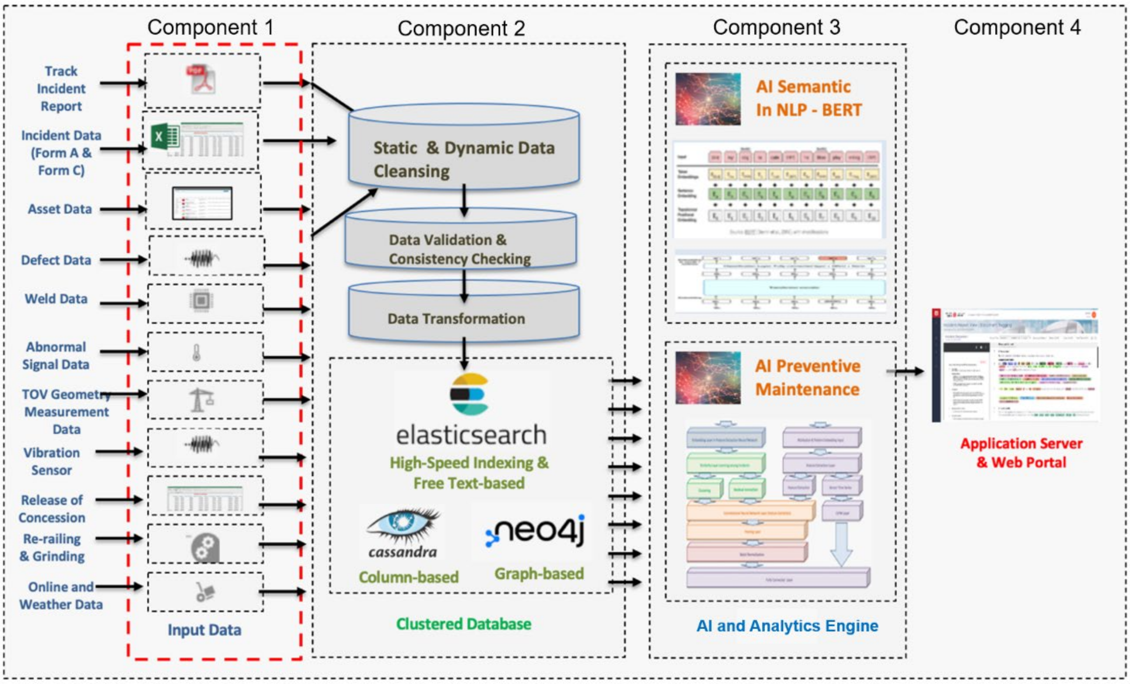

Figure 1: System Architecture

3.1 Component 1: Input Data (Ingestion Module)

Data is ingested into the system from different data sources, including static data (asset data, incident reports and maintenance records), dynamic data (real-time conditions) and online data (weather and news), and listed in Figure 1. The data is inserted into the data lake of the schema designed.

3.2 Component 2: Clustered Database

The Clustered Database is formed by integrating the high-speed indexing engine DB for free text, Columnized DB for real-time data and Graph-based DB for incident correlations for Pway data and key-value identified.

High-Speed Indexing Engine - Elastic-Search: Application Data can be searched at high-speed using textual indexing search. Based on this, we can efficiently retrieve and analyze the relationship between incidents and various data sources.

Columnar-based Database – Cassandra: Real-time sensor data is stored in Cassandra with elastic-search to speed up the retrieval process.

Graph-based Database - Neo-4J: We store relationships of entities and incidents in a graph-based database. This kind of data is transformed into graph-based structures to emulate the fundamental relationship of the incident and equipment for better queries by AI for analysis.

3.3 Component 3: AI and Analytics Engine

The system can feed data in the AI and analytics engine to process different data sets, data transformation and calculate the likelihood of permanent way incidence occurrence.

3.3.1 Semantic AI Model - Q&A Module

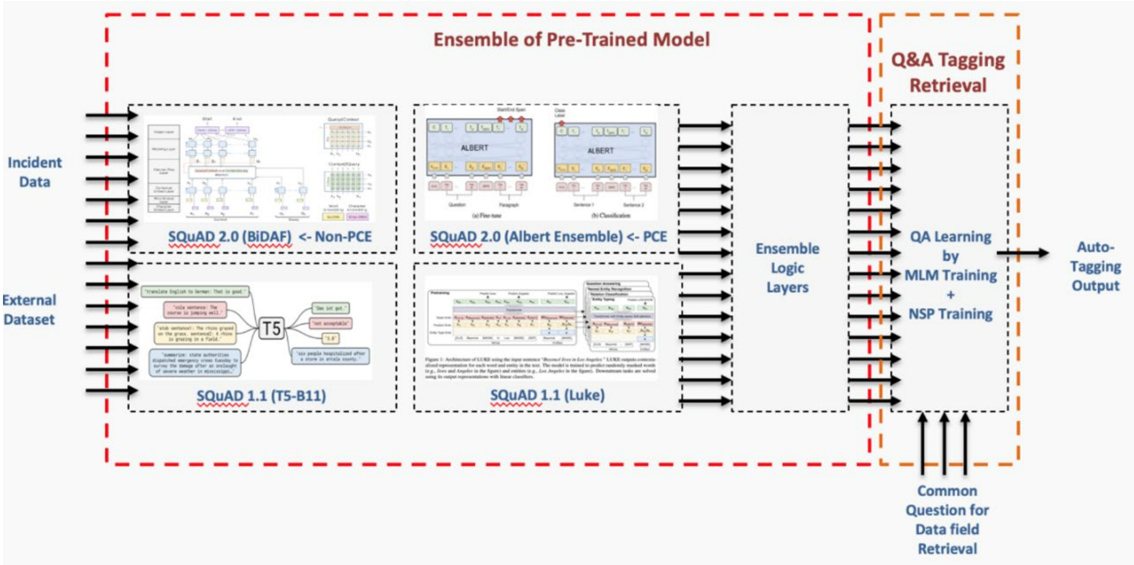

Figure 2: Semantic AI Model

The semantic AI model consists of two components as follows: -

Ensemble of Pre-Trained Models

Different pre-trained models have strengths in different document structures and sentence patterns. In Phase 1, we mix several models together to obtain a more powerful ensembled model. The final model is an all-model-ensemble that involves Albert-Ensemble Model and several BERT-variant models. In Phase 1, we fine-tune the model with incident data. In Phase 2, we further fine-tune the model with operating railway data on concession application data.

Q&A Tagging Retrieval

This part is cast as a supervised learning task. The methodology is similar to Masked Language Model (MLM) and Next Sentence Prediction (NSP) (Devlin et al., 2018), which is summarized below.

Step 1: Determine the Existence of the Answer in the Sentence

Similar to Phase 1, we utilize AI QA to query the data. The data may exist or may not exist. We use SQuAD 2.0-variant model to determine if an answer exists. If the answer does not exist in the data, the models spot the non-existence of the answer. In our previous study, out of 100 paragraphs and 10 questions, 60% of answers do not exist. Therefore, detection of the existence of an answer enhances the resultant accuracy significantly.

Step 2a: Retrieving "Tag" by using AI QA Model
After deciding the existence by the above model, we switch to an all-model-ensemble (Last Model in Phase 1) model with higher accuracy to find the answer to the question. To retrieve the keyword "Tag" in the best general sense, our tuned model can retrieve some general sense information, such as "What is the last inspection date?" Since dates are general terminology that can be understood by T5 model (Raffel et al., 2020), the T5 model is likely to return the result even with the pre-trained model only.

Step 2b: Model Better in Specific Data – Special Training on Release of Concession (RoC) Data
If the answer involves specific data keywords from data, special data model with training dataset is required. We employ the XLM Roberta model (Liu et al., 2019) to return the answer with keywords. For example, given a question, "What is the cause of the incident?" keywords such as "Broken Rail" are quite specific to data. Blending the tagged keywords from documents can calculate the relevance of the keywords. The model can return a list of answers in descending order of relevance.

Step 3 Answers from both models are mixed and output as a final result
Answers are feedback to find a better fit in different situations of text.

3.3.2 AI Model for Predictive Maintenance in a Pway System

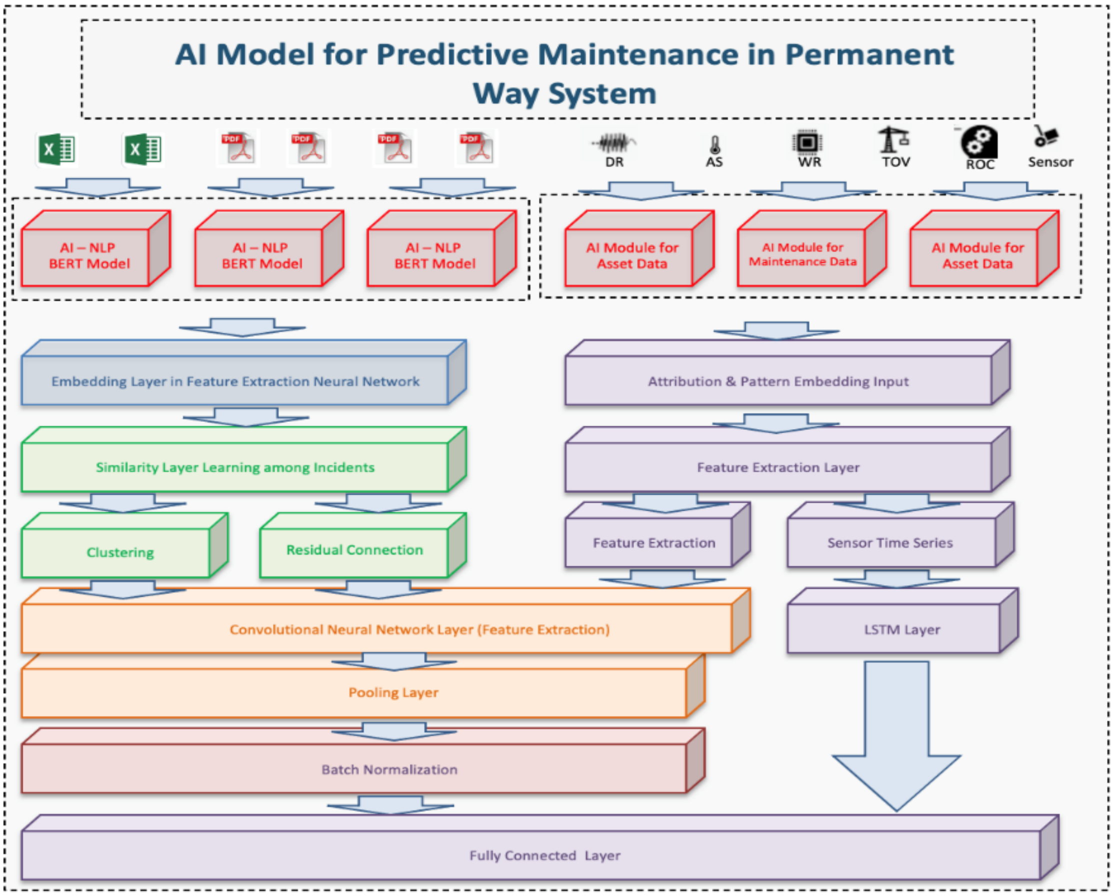

Figure 3: Model Structure

<u>Step 1: Use of Proprietary ALBERT-Ensemble Model</u>

Proprietary ALBERT-Ensemble Invariant Model, the State-of-the-Art NLP AI model created in Phase 1, retrieves different attributes from different incidents in the report with key-value pairs. The key-value pairs contain essential information from the incident.

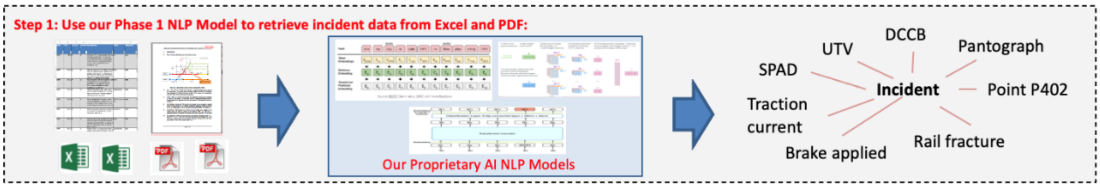

Figure 4: retrieve incident data from data source with NLP model

Our proprietary model uses the encoder-decoder architecture and is trained with Railway Incident Data.

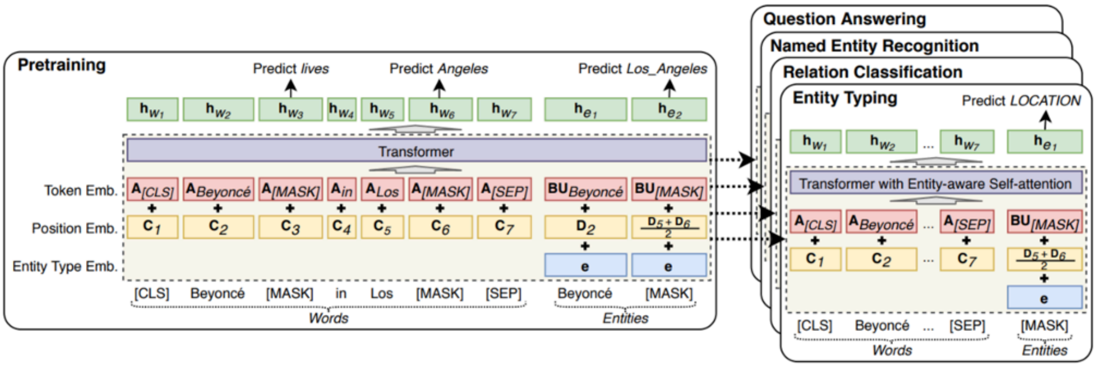

Figure 5: Pre-Trained Model

We have successfully retrieved different railway-specific attributes using named entities recognition. Our natural language processing modules can successfully retrieve textual information from textual descriptions in PDF and Excel files. Observation sets of the railway incidents are constructed with attributes extracted from Pway data and are stored in a high-speed big data elastic-search data lake. Our proposed algorithms are SQuAD2.0 compliant with state-of-the-art benchmarking with Railway data adaptation to achieve the best textual AI performance.

<u>Step 2: Use of Embedding</u>

We utilize EMBEDDING Layer techniques to compress incident attributes into condensed N-Dimensional Tensor matrices for subsequent layer input.

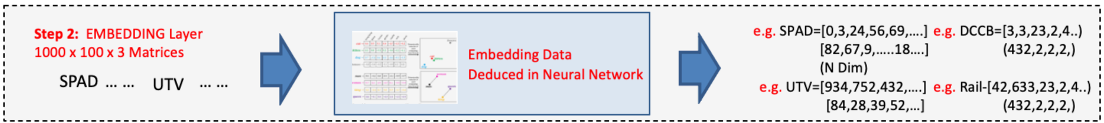

Figure 6: Illustration of using embedding layer

The embedding techniques translate all textual semantic meanings into matrices (i.e. mathematical representations), and orthogonal feature selection techniques compress these mathematical representations. In layman's terms, semantic meanings are digitized to ease the vector calculations of AI classifiers.

Step 3: Use of Similarity-Embedding

This step correlates different incidents by different selected key-value pairs and attributes. Apply the machine learning in learning the similarity of incidents using previous embedding data.

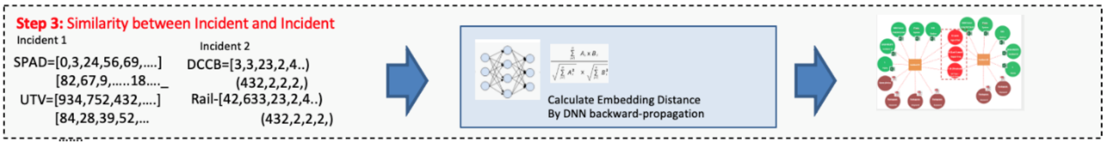

Figure 7: similarity calculation

The embedding matrices from step 2 shall be used to calculate the similarity matrices. The operation simulates N-dimensional dot-products of features. This step optimizes the similarity measures after the shuffling of all combinations of inputted feature dimensions. Special techniques such as Minkowski filtering may be enforced for regularization to prevent the overfitting of sparse data from incident reports.

Step 4: Use of Hybrid Decision Tree

We use both hybrid-tree-based decision-tree clustering and residual connection in this step. The tree-based decision tree clustering provides a clustering by using information-gain-index. It compensates for the problem of lack of data.

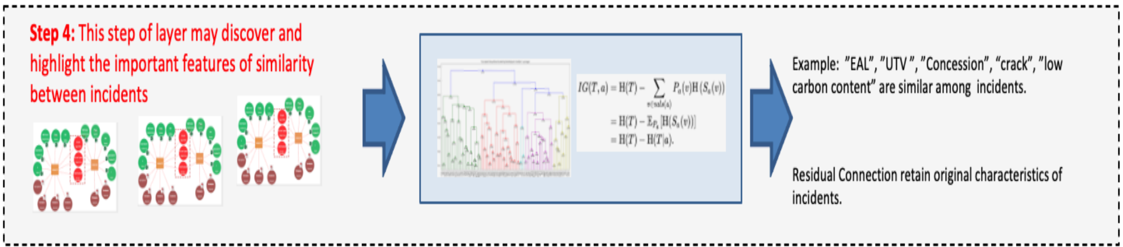

Figure 8: Illustration of using hybrid decision tree

Use of Residual Connections: Similar to the famous RESNET, we use residual connections to preserve the identity data and blend it with hybrid clustering tree output.
Resemblance of Explainable AI (X-AI): The X-AI techniques explain the attributions of contributing factors. It is similar to the random forest concept. The perturbation of parameters in attributing factors in feature trees prevents optimizing the best hyper-parameters while maintaining human-readability advantages.

Step 5: Utilizing Convolutional Neural Network Layers for Feature Extraction

AI network retrieves important combinations of behavioral patterns among primitive features. The pooling layer aggregates the results from high dimensions to lower dimensions. In layman's terms, it summarizes the pattern result by aggregation.

Step 6: Feature Extraction with Patterns

We extract features from several data sources for prediction. The data exhibits patterns with a combination of attribute occurrences. For example, Track Incident Records (TIR) correlate with welding and Abnormal Signal (AS) records deeply. The data records, such as the chainage with frequent welding records, indicate a higher likelihood of an incident. The in-depth features, like the attribute of vibration sensor data, could be extracted in a time-series pattern. The convolution layers extract relationships from these features and the combined patterns to feed the final regression layer for decision-making.

Step 7: Long Short-Term Memory (LSTM) Layers
These layers retrieve the feature of incidence across time. For example, the trends in vibration sensor and ultrasonic sensor values may be a critical factor in predicting the probability of an incident. It retrieves the feature and characteristics in time series data. It will be fed to the final regression layer to mix the final decision.

Step 8: Final Decision-Making Layers
We utilize fully connected layers to act as N-dimensional universal function approximators. They output a single continuous value as the likelihood of incident occurrence at a particular chainage through all the convolution output features criteria and weightings.

3.4 Component 4: Application Server (Web Portal)
Application server with web portal is built to display the AI model outcomes on dashboards and knowledge graph.

RESULTS

4.1 Ranking and Contributing Factors of Railway Track Systems
In the early phase of the long-term analysis journey, it is expected that it is not known which and what pieces of data are essential in deriving the results. Therefore, the subset of data is selected in this analysis to balance the data collection efforts of the operation team. Hence, the goal here is to figure out more potential insights. Those insights shall assist in better selecting and collecting data for future analysis and model building. This subsection has outlined a list of factors that may have contributed to the incidents.

Correlations, ranking and contributing factors of railway track incidents were identified and showed in the result of semantic models. 119 contributing factors are identified by the semantic AI models based on operating railway training data from Jan 2016 to Dec 2021, and these factors are of varying probability to contribute to TIR incidents. Table 1a shows the contributing factors identified by semantic models and probability of contribution to TIR incidents. Track geometry, small track curve radius, abnormal ultrasonic reading in WR record and switch crossing are most significant contributing factors.

Table 1b shows vibration sensors related contributing factors as identified by the semantic AI models and probability of contribution to TIR incidents. Those vibration sensors such as "Spring Deflection", "Track CrossLevel" and "Lateral Body Acceleration" are examples of factors that are not easily noticeable but have higher probability of contribution to TIR incidents. They give valuable insights to maintainer on their predictive maintenance planning.

Contributing Factor Group	ContributingFactor	Probability(TIR	Description
Geometry (AS)	AS Radius < 500	100%	Track radius < 500mm in AS record
Geometry (AS)	AS Radius < 400	100%	Track radius < 400mm in AS record
Geometry (AS)	AS Radius < 300	100%	Track radius < 300mm in AS record
Geometry (AS)	AS Sharp Curve High Leg	93%	SCHL track
Geometry (WR)	WR Radius < 300	88%	Track radius < 300mm in WR record
WR Abnormal Reading	WR Offside Abnormal Value	86%	Abnormal ultrasonic reading (Offside) in WR record
Geometry (AS)	AS Crossing	77%	Switch crossing in AS record
AS Abnormal Signal	AS Squat	75%	Squat defect in AS record
Geometry (WR)	WR Radius < 400	73%	Track radius < 400mm in WR record
AS Abnormal Signal	AS Shelling	72%	Shelling defect in AS record
Location-Related	Chainage	72%	Location in chaingage
Location-Related	Up Track / Down Track	72%	Track direction
Location-Related	Left-hand rail / Right-hand rail	72%	Rail hand
Location-Related	Station	72%	Location from/to Station
Time-Related	Time-Year Trend For Station	72%	Year of track incident
Time-Related	Time-Month	72%	Month of track incident
Geometry (WR)	WR Radius < 500	70%	Track radius < 500mm in WR record
AS Abnormal Signal	AS Crack	70%	Crack defect in AS record
Geometry (WR)	WR Sharp Curve High Leg	68%	SCHL track in WR record
WR Abnormal Reading	WR Degree 38 Abnormal Valu	66%	Abnormal ultrasonic reading (38 Deg probe) in WR record
WR Abnormal Reading	WR Degree 70 Abnormal Valu	61%	Abnormal ultrasonic reading (70 Deg probe) in WR record
WR Abnormal Reading	WR Onside Abnormal Value	60%	Abnormal ultrasonic reading (Onside) in WR record
WR Abnormal Reading	WR Degree 45 Abnormal Valu	50%	Abnormal ultrasonic reading (45 Deg probe) in WR record
Vibration Sensor Abnormal Reading	VIB Lp1 > 5	50%	Spring Deflection 1 >5g
Geometry (WR)	WR Crossing	49%	Switch crossing in WR record
Recent Service/Recent Environment Changes (Weat	Temperature Difference > 5	32%	Daily temperature difference > 5 Deg within 5 days
Recent Service/Recent Environment Changes (Weat	Temperature Difference > 4	31%	Daily temperature difference > 4 Deg within 5 days
Recent Service/Recent Environment Changes (Weat	High Temperature	31%	High Temperature
Recent Service/Recent Environment Changes (Weat	Temperature Difference > 2	30%	Daily temperature difference > 2 Deg within 5 days
Recent Service/Recent Environment Changes (Weat	Temperature Difference > 3	29%	Daily temperature difference > 3 Deg within 5 days
Vibration Sensor Abnormal Reading	VIB Crosslevel > 10	29%	Track CrossLevel>10mm
Vibration Sensor Abnormal Reading	VIB Bod2 Acc y > 0.1	26%	Lateral Body Acceleration (Driver Side)>0.1g
Vibration Sensor Abnormal Reading	VIB Acc2 > 8	25%	Vertical Acceleration 2>8g
Vibration Sensor Abnormal Reading	VIB Acc3 > 20	25%	Vertical Acceleration 3>20g
Vibration Sensor Abnormal Reading	VIB Profile z1 > 10.8	25%	Vertical Rail Profile 1>10.8mm

Table 1a: Contributing factors identified by semantic models and probability of contribution to TIR incidents (only contributing factors with probability >25% are shown)

Description	ContributingFactor	Probability(TIR)
Spring Deflection 1 > 5mm	VIB Lp1 > 5	50%
Track CrossLevel > 10mm	VIB Crosslevel > 10	29%
Lateral Body Acceleration (Driver Side) > 0.1g	VIB Bod2 Acc y > 0.1	26%
Vertical Acceleration 2 >8g	VIB Acc2 > 8	25%
Vertical Acceleration 3 >20g	VIB Acc3 > 20	25%
Vertical Rail Profile 1 >10.8mm	VIB Profile z1 > 10.8	25%
Lateral Body Acceleration (Gangway) > 0.1g	VIB Bod1 Acc y > 0.1	24%
Bogie Bounce Front > 2mm	VIB Bounce Fr > 2	24%
Bogie Bounce All > 2mm	VIB Bounce > 2	22%
Vertical Acceleration 4 >15g	VIB Acc4 > 15	21%
Vertical Acceleration 1 >8g	VIB Acc1 > 8	21%
Vertical Acceleration All >10g	VIB Acc > 10	20%
Vertical Rail Profile 4 >10.8mm	VIB Profile z4 > 10.8	19%
Bogie Bounce Rear > 2mm	VIB Bounce Rr > 2	17%
Spring Deflection All >3mm	VIB Snd > 3	17%
Ride Comfort Rear Vert >0.18m/s2	VIB Cc 2z > 0.18	17%
Spring Deflection 3 >3mm	VIB Lp3 > 3	17%
Spring Deflection 4 >3mm	VIB Lp4 > 3	16%
Bogie Rock Front > 5mm	VIB Rock Rr > 5	15%
Spring Deflection 2 >2mm	VIB Lp2 > 2	15%
Ride Jerk Rear Lat > 0.3g/s	VIB Bod2 Lat Jerk > 0.3	13%
Bogie Rock Front >5mm	VIB Rock Fr > 5	13%
Bogie Rock All >5mm	VIB Rock > 5	11%
Ride Jerk Front Lon > 0.2g/s	VIB Bod1 Lon Jerk > 0.2	9%
Ride Jerk Rear Lon > 0.2g/s	VIB Bod2 Lon Jerk > 0.2	9%
Ride Jerk Long > 0.2 g/s	VIB Lon Jerk > 0.2	9%
Ride Jerk Lat > 0.5 g/s	VIB Lat Jerk > 0.5	8%
Ride Comfort Front Vert > 0.2m/s2	VIB Cc 1z > 0.2	7%
Ride Comfort Rear Lat >0.15m/s2	VIB Cc 2y > 0.15	7%
Ride Comfort Rear Long > 0.08m/s2	VIB Cc 2x > 0.08	5%
Ride Comfort Front Long >0.8m/s2	VIB Cc 1x > 0.08	5%
Ride Comfort Front Lat >0.2m/s2	VIB Cc 1y > 0.2	3%

Table 1b: Contributing factors (vibration sensors) identified by semantic models and probability of contribution to TIR incidents

***Case Study*:** We analyze the contributing factors (CFs) for incidents through a knowledge graph based on the operating railway training data from Jan 2016 to Dec 2021 and give an example of how to find out the CFs to the incident PWS-818-20. Figure 9 displays a knowledge graph illustrating relationships between factors and the incident. The strength of the edges in the graph indicates how strong the connection between contributing factors and TIR. Figure 10 shows the contributing factors summary and ranking for all TIR incidents. Many panels list the abnormal records related to the incident, and the Chainage Probability panel also provides the incident probability along chainage.

Figure 11 shows the vibration sensor record graph that can display a particular type of vibration sensor along the chainage of a particular date. Various thresholds are incorporated on the graph to identify vibration sensor value spikes.

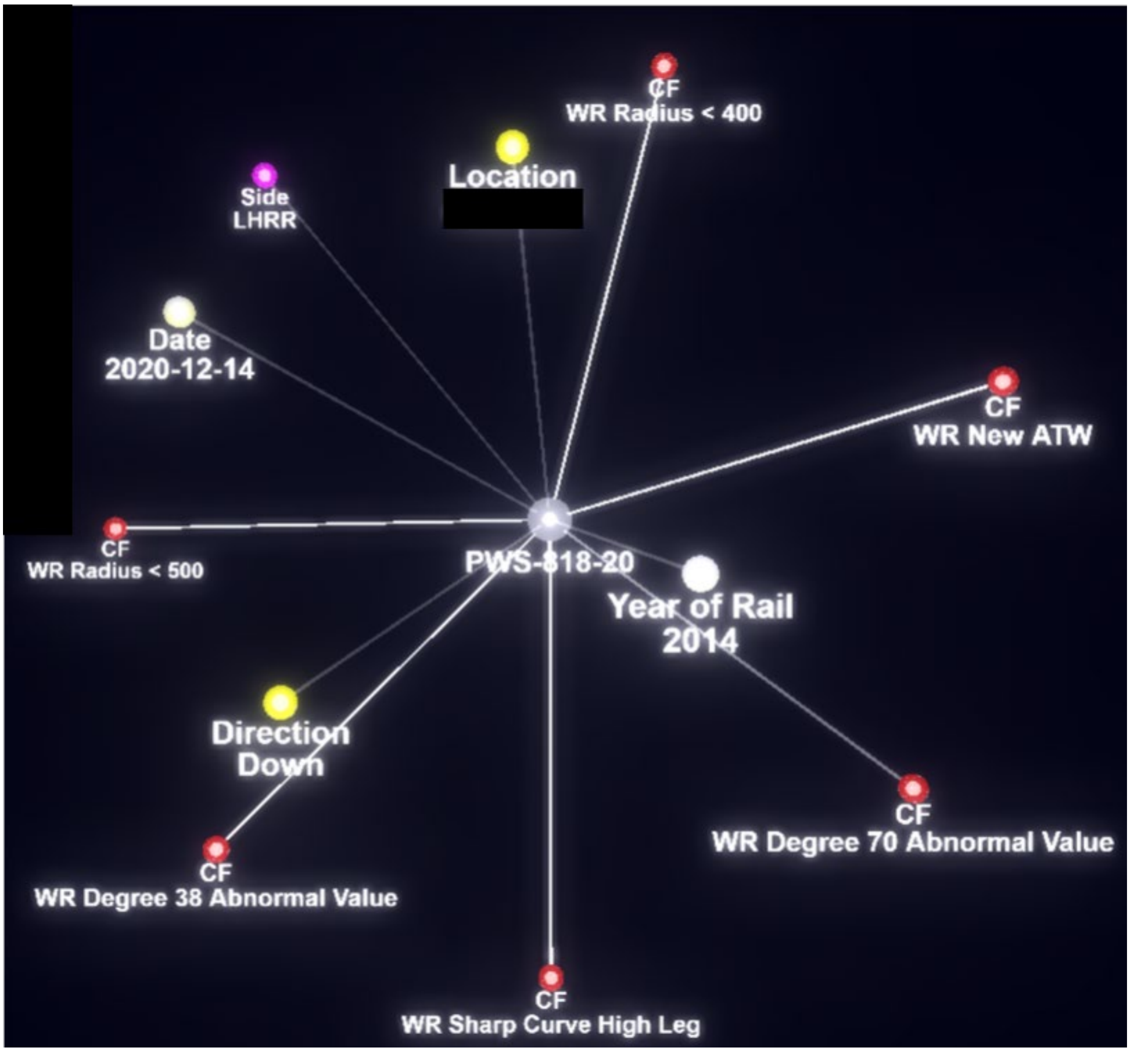

Figure 9: Example of identifying contributing factors for an incident (PWS-818-20) through the knowledge graph

Contributing Factor (CF) Summary Search

Rank	Category	Contributing Factor	Description	Occurence	Probability (TIR) ▼
1	Geometry (AS)	AS Radius < 500		25	100%
2	Geometry (AS)	AS Radius < 400		18	100%
3	Geometry (AS)	AS Radius < 300		11	100%
4	Geometry (AS)	AS Sharp Curve High Leg		19	93%
5	Geometry (WR)	WR Radius < 300		22	88%
6	WR Abnormal Reading	WR Offside Abnormal Value		18	86%
7	Geometry (AS)	AS Crossing		44	77%
8	AS Abnormal Signal	AS Squat		151	75%
9	Geometry (WR)	WR Radius < 400		35	73%
10	AS Abnormal Signal	AS Shelling		119	72%
11	Location-Related	Chainage		883	72%
12	Location-Related	Up Track / Down Track		883	72%
13	Location-Related	Left-hand rail / Right-hand rail		883	72%
14	Location-Related	Station		883	72%
15	Time-Related	Time-Year Trend For Station		883	72%
16	Time-Related	Time-Month		883	72%
17	Geometry (WR)	WR Radius < 500		45	70%
18	AS Abnormal Signal	AS Crack		117	70%
19	Geometry (WR)	WR Sharp Curve High Leg		45	68%
20	WR Abnormal Reading	WR Degree 38 Abnormal Value		58	66%
21	WR Abnormal Reading	WR Degree 70 Abnormal Value		68	61%
22	WR Abnormal Reading	WR Onside Abnormal Value		15	60%
23	WR Abnormal Reading	WR Degree 45 Abnormal Value		2	50%
24	Vibration Sensor Abnormal Reading	VIB Lp1 > 5	Spring Deflection 1 >5g	3	50%
25	Geometry (WR)	WR Crossing		34	49%
26	Recent Service/Recent Environment Changes (Weather)	Temperature Difference > 5		106	32%
27	Recent Service/Recent Environment Changes (Weather)	Temperature Difference > 4		191	31%
28	Recent Service/Recent Environment Changes (Weather)	High Temperature		84	31%
29	Recent Service/Recent Environment Changes (Weather)	Temperature Difference > 2		555	30%
30	Recent Service/Recent Environment Changes (Weather)	Temperature Difference > 3		337	29%
31	Vibration Sensor Abnormal Reading	VIB Crosslevel > 10	Track CrossLevel>10mm	7	29%
32	Vibration Sensor Abnormal Reading	VIB Bod2 Acc y > 0.1	Lateral Body Acceleration (Driver Side)>0.1g	76	26%
33	Vibration Sensor Abnormal Reading	VIB Acc2 > 8	Vertical Acceleration 2>8g	64	25%
34	Vibration Sensor Abnormal Reading	VIB Acc3 > 20	Vertical Acceleration 3>20g	16	25%
35	Vibration Sensor Abnormal Reading	VIB Profile z1 > 10.8	Vertical Rail Profile 1>10.8mm	19	25%
36	Vibration Sensor Abnormal Reading	VIB Acc2 > 10	Vertical Acceleration 2>10g	54	24%
37	Vibration Sensor Abnormal Reading	VIB Bod1 Acc y > 0.1	Lateral Body Acceleration (Gangway)>0.5g	79	24%
38	Vibration Sensor Abnormal Reading	VIB Bounce Fr > 2	Bogie Bounce Front >2mm	20	24%
39	Vibration Sensor Abnormal Reading	VIB Acc2 > 5	Vertical Acceleration 2>5g	98	23%
40	Vibration Sensor Abnormal Reading	VIB Bounce Fr > 1.5	Bogie Bounce Front >1.5mm	50	23%
41	Vibration Sensor Abnormal Reading	VIB Acc3 > 10	Vertical Acceleration 3>10g	47	23%
42	Recent Service/Recent Environment Changes (Weather)	Low Temperature		6	22%

Figure 10: Contributing factors summary and ranking for all TIR incidents

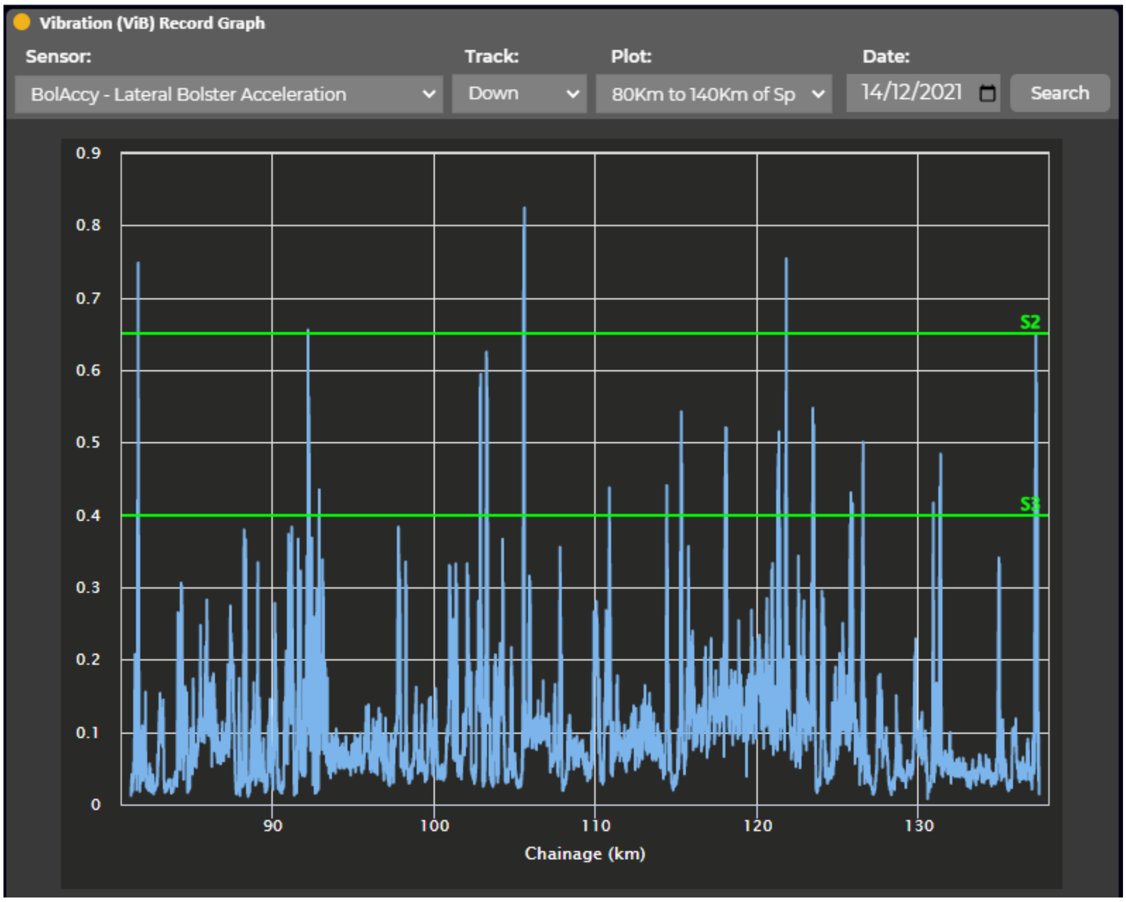

Figure 11: Vibration sensor record graph

4.2 Incident Probability Prediction

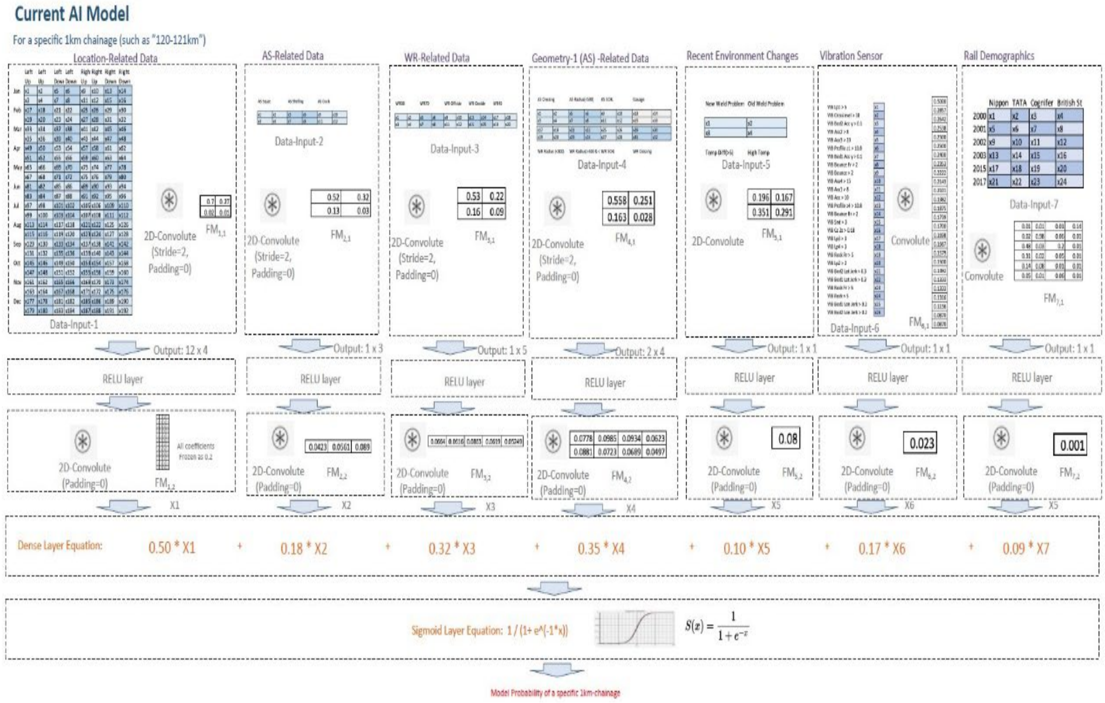

Figure 12: Overview of the Railway AI Predictive Maintenance Model

In order to make a long-term development, we establish a railway framework. The framework actually demonstrates the use of different categories of Pway data sources ranging in location-related data, AS-related data, WR-related data, track geometry-related data, environment changes data, vibration sensor data and rail demographics to find out the correlation with TIR for predicting railway track incidents along each track section (1 km chainage). Figure 12 shows the overview of our Railway AI Predictive Maintenance Model.

Table 2 shows the result of predictive incident probability along the chainage by the AI model trained on operating railway training data from Jan 2016 to Aug 2021. Besides the incident probability, the AI model also indicates on what basis the incident probabilities were calculated.

Chainage	Direction	Reason	Probability (%)
1	Downtrack	WR R70 Reading, SCHL	80
2	Uptrack	WR R70 Reading, SCHL	78
3	Downtrack	WR R38 Reading	85
4	Uptrack	WR R70 Reading, SCHL	56
5	Uptrack	Sharp-Curve-High Leg	83
6	Uptrack	SCHL, WR Reading R38	86
7	Uptrack	WR R70 Reading, SCHL	76
8	Downtrack	Sharp-Curve-High Leg	52
9	Uptrack	WR R70 Reading, SCHL	57
10	Downtrack	Sharp-Curve-High Leg	63

11	Uptrack	WR R70 Reading, SCHL	62
12	Uptrack	WR R70 Reading, SCHL	61
13	Downtrack	WR R70 Reading, SCHL	61

Table 2: Predicted Probability at Specific Locations

4.2.1 Accuracy

We test the AI model trained on TIR cases from Jan 2016 to Aug 2021 with the recent data of TIR cases from Sep 2021 to Apr 2022. Table 3 indicates the model can achieve an accuracy of at least 53% and a peak at 69% under different probability thresholds.

Probability Threshold	Correct Predictions	**Accuracy**	False Positives	False Negatives
0.1	62	**53%**	10%	36%
0.2	70	**60%**	14%	26%
0.3	76	**66%**	16%	19%
0.4	79	**68%**	17%	15%
0.5	78	**67%**	19%	14%
0.6	80	**69%**	21%	10%
0.7	80	**69%**	23%	8%
0.8	75	**65%**	29%	6%
0.9	76	**66%**	33%	2%

Table 3: Results of testing the Railway AI Predictive Maintenance Model (TIR cases from Jan 2016 to Aug 2021) with current data (TIR cases from Sep 2021 to Apr 2022)

The AI model is re-trained on TIR cases from Jan 2016 to Dec 2021 and tested with the most recent data of TIR cases from Jan 2022 to Jun 2022. Table 4 indicates the model can achieve an accuracy of 72% under probability thresholds of 0.4 to 0.7.

Probability Threshold	Correct Predictions	**Accuracy**	False Positives	False Negatives
0.1	69	**59%**	36%	4%
0.2	73	**63%**	27%	10%
0.3	77	**66%**	20%	14%
0.4	83	**72%**	15%	14%
0.5	83	**72%**	9%	19%
0.6	83	**72%**	6%	22%
0.7	83	**72%**	4%	24%
0.8	85	**73%**	3%	24%
0.9	87	**75%**	0%	25%

Table 4: Results of testing the Railway AI Predictive Maintenance Model (TIR cases from Jan 2016 to Dec 2021) with current data of TIR cases from Jan 2022 to Jun 2022

In order to test the incident prediction ability of the AI model on other railway line, we applied another operating railway dataset (from Apr 2021 to Dec 2021) and tested the AI model with another recent set of data of TIR cases from Jan 2022 to Apr 2022. Table 5 indicates the AI model can achieve an accuracy of at least 52% and a peak at 82% under different probability thresholds.

Probability Threshold	Correct Predictions	**Accuracy**	False Positives	False Negatives
0.1	39	**63%**	32%	5%
0.2	39	**63%**	32%	5%
0.3	51	**82%**	13%	5%
0.4	51	**82%**	13%	5%
0.5	45	**73%**	8%	19%
0.6	35	**72%**	3%	40%
0.7	34	**56%**	0%	45%
0.8	33	**55%**	0%	47%
0.9	32	**52%**	0%	48%

Table 5: Results of testing the Railway AI Predictive Maintenance Model on other set of operating railway data (TIR cases from Apr 2021 to Dec 2021) with current data of TIR cases from Jan 2022 to Apr 2022

One of our biggest challenges of this AI project is that the number of incidents is small due to the effective maintenance regime of MTR over the years. The implication that we derived from the current dataset would require more incident data to calibrate the significance and relative weightings in the prediction. Despite of this challenge, it is very encouraging to start on the long-term analytic journey.

4.2.2 Visualization

The incident probability for the chainage as derived by the AI model is given and classified as "High", "Medium", and "Low". "Red" (High), "Yellow" (Medium), and "Green" (Low) probabilities are shown on the dashboard. With the help of the AI model, we could quickly identify the potentially higher-risk equipment and areas in railway systems and consider appropriate maintenance attention.

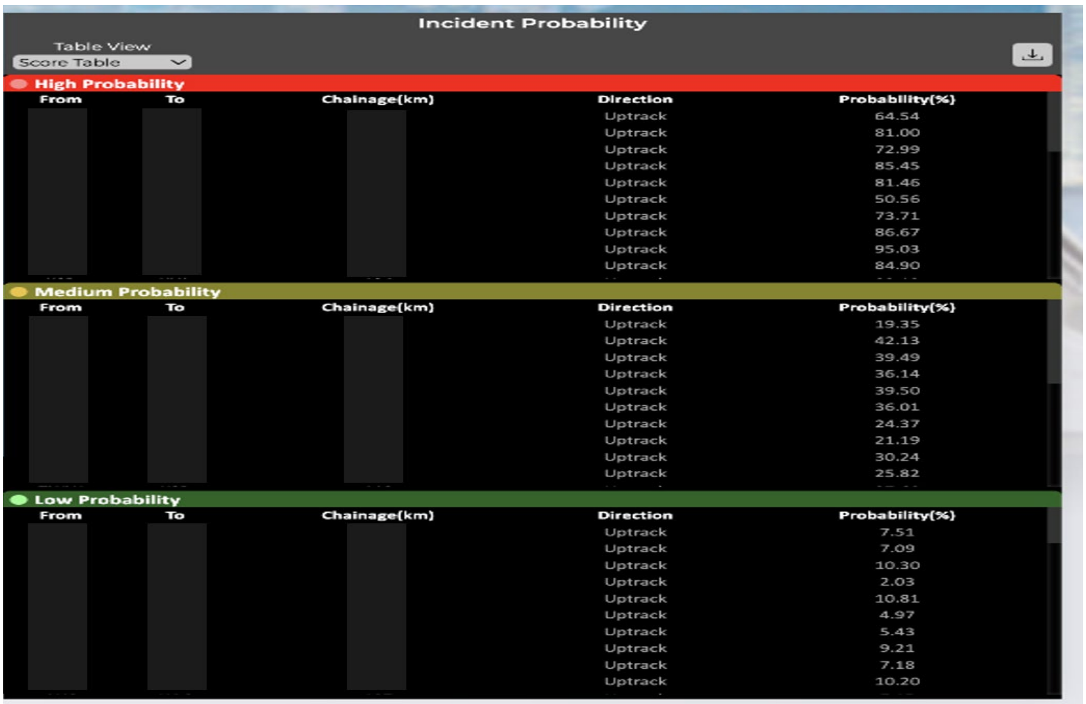

Figure 13: Incident probability along chainage

CONCLUSION

A pilot project has been successfully implemented for developing a semantic AI model with "Railway Schema", "Semantic Model" and "Railway AI Predictive Maintenance Model" that aims to identify risks and predict potential incidents in the railway track system.

In this study, we integrate various data sources of railway track incidents data into a clustered database. The data is stored as structured knowledge, enabling powerful and efficient analytical queries. Based on the clustered database, we develop a Q&A module within the AI model that allows analysts to efficiently and effectively discover knowledge about railway incidents from the data. To predict the likelihood of railway track incidents and identify equipment and areas at risk, we further propose an AI model incorporating the semantic features extracted from the raw data. Our case study of contributing factors, incident probability prediction and visualization of potential risk demonstrates the effectiveness of the system based on the available dataset in the study.

By exploiting semantic AI technologies, we can integrate information from various data sources and apply computational tools to draw comprehensive insights of track incidents. A novel predictive maintenance approach using semantic AI technology was developed to determine the ranking and contributing factors (including those unnoticeable features) of railway track incidents in the coming 12 months to empower the maintainers with predictive early warnings, historical case matching, and actionable intelligence. It will bring asset management, data storage, data searching and ultimately predictive maintenance analysis for railway track into a new era by use of a semantic AI model. It has demonstrated an innovative way to enable predictive maintenance by AI, supplementing the traditional maintenance method in the railway industry.

This project is a very encouraging starting point of the long-term analytic journey while it was noted that prediction accuracy would need to be further optimized as there are rare past incident data. The implication that we derived from the current dataset probably requires more incident data to calibrate the significance and relative weightings in the prediction. More incident data may be obtained, e.g. further historical data, to the AI model for learning.

Our future plan is to explore this application in other railway systems, such as the signalling, power and rolling stock, for devising predictive maintenance to contribute further enhancement of the overall safety and reliability of the railway system in Hong Kong and in the industry.

Keywords: Semantic AI; Predictive Maintenance; Pway

REFERENCES

Chaudhri, V., Baru, C., Chittar, N., Dong, X., Genesereth, M., Hendler, J., Kalyanpur, A. ., Lenat, D. ., Sequeda, J. ., Vrandečić, D., & Wang, K. (2022). Knowledge Graphs: Introduction, History, and, Perspectives. *AI Magazine*, 43(1), 17-29. https://doi.org/10.1609/aimag.v43i1.19119

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *ArXiv Preprint ArXiv:1810.04805*.

Ho, T. K. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, *1*, 278–282.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. *ArXiv Preprint ArXiv:1607.01759*.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *ArXiv Preprint ArXiv:1907.11692*.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137.

MacQueen, J. (1967). Classification and analysis of multivariate observations. *5th Berkeley Symp. Math. Statist. Probability*, 281–297.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, *26*.

paperswithcode. (n.d.). Question Answering on SQuAD2.0. *Question Answering on SQuAD2.0*. https://paperswithcode.com/sota/question-answering-on-squad20

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (1802). Deep contextualized word representations. CoRR abs/1802.05365 (2018). *ArXiv Preprint ArXiv:1802.05365*.

Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., & others. (2018). *Improving language understanding by generative pre-training*. OpenAI.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P. J., & others. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, *21*(140), 1–67.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, \Lukasz, & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in Neural Information Processing Systems*, *32*.

Zhang, Y., & Wallace, B. (2015). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. *ArXiv Preprint ArXiv:1510.03820*.