1. Project content

In China, there are more than 6 million base stations, which brings huge capital investment including construction cost, operating cost and maintenance cost. China Mobile's total revenue in 2019 was 700 billion, of which one third was used for base station operation and maintenance. In 2021, China will have a 400 billion operation and maintenance market, and the global market will be nearly 1 trillion. So BS investment is increasing year by year.

When a network failure occurs at a base station, two types of alarms will be generated: minor alarms and out-of-service alarms. The traditional operation and maintenance mode is to repair the base station after failure, which is relatively passive. Our innovative model is to build an AI model and use the secondary alarms generated by the base station to predict future out-of-service alarms, realize base station early warning, maintain base stations in advance, avoid service interruption, and improve user communication quality.

2. Solution

The original data is complex and disordered. And we should transform it to train data which is succinct and clear. It contains ID column, feature columns and label column. The feature represents training samples and the label is the prediction target.

✓ "Data is the base, features are king", so start with data processing.

What is provided is the complex and disorderly original data of "alarm occurrence time", which needs to be transformed into the concise and clear training data of "ID, feature, label".

✓ The model adopts GBDT model:

Decision tree: The base model is a tree model.
Gradient: Use the negative gradient of the loss function as the residual to fit.
Improvement: The original model continuously linearly superimposes the residual prediction value to make the decision tree more accurate.

✓ Model optimization is mainly in three aspects:

5-fold cross-validation prevents over-fitting; grid search method finds the optimal algorithm parameters; Heamy model fusion, through the complementary advantages of sub-models, superimposed to generate the optimal model.

3. Innovation

✓ Innovation point 1: Shift to generate labels.

Filling in the missing dates and shifting the time upwards can efficiently and accurately generate training set labels to achieve accurate prediction. Experiments have proved that due to the high label accuracy, this step becomes the fastest solution for improving scores.

✓ Innovation point 2: Threshold adjustment

The ratio of label 0 to label 1 is about 5:1, and the data set is not balanced. Traditional oversampling and undersampling are not effective due to the drawbacks of data duplication and missing. We adopt the threshold adjustment based on the 1/0 ratio of 0.165 as the benchmark value to effectively improve the prediction accuracy.

✓ Innovation point 3: feature engineering

(a) Basic features: the number of occurrences of each type of alarm
(b) Derivative features: 300-dimensional features such as the statistical features of alarms in the previous N days, the statistical features of de-service alarms, and the statistical features of (8 types) alarm types. "Feature is king", multi-dimensional detailed feature engineering ensures the accuracy and generalization ability of the model.

4. Project advantages

✓ Data without labels is of no value to machine learning, and our solution can automatically and accurately generate labels, effectively reducing labor costs.

✓ Our prediction accuracy is high. Final F1-score is 74.6667%, 1.4% higher than the second place.

✓ The training speed is fast, the model training only takes 3-5 minutes, which is much faster than neural network models such as LSTM

✓ Strong generalization ability, collecting two-year operating data of tens of thousands of base stations to train the model, so that the model has a strong generalization ability

✓ Strong scalability. When accessing information such as base station power and power consumption, data processing and feature construction can also be input into the model

At present, most of the project plans are to use 5G communication services, and our plan is committed to making 5G communications intelligent, efficient, and low-cost, so that 5G communications can better serve everyone.

5. Future Market

At present, 1/3 of China Mobile's total revenue is used for base station operation and maintenance. The cost of base station operation and maintenance is increasing year by year, and the market scale is very large. Among them, the traditional operation and maintenance mode, as the equipment ages over time, the cost will increase; the AI predicts the operation and maintenance mode, the early investment predicts the cost, avoids the aging failure of the equipment, and the operation cost is low and stable.

According to the global predictive maintenance market trend, in 2022, the annual growth rate of the global predictive maintenance market will exceed 27%. The development of artificial intelligence has saved billions of operation and maintenance costs, and the number of suppliers providing predictive maintenance services has doubled. And it involves many key areas such as big data, artificial intelligence, and 5G communications. This is the first time AI has been applied to the field of 5G base station operation and maintenance. We hope to create a new 5G smart network operation and maintenance model.