Motivation
As telecommunication companies seek to improve network coverage and maintain customer satisfaction, they need to constantly monitor customer complaints and the status of their networks.

This project, in collaboration with Tupl Inc. seeks to reduce the latency between the time a network coverage area experiences issues and when the network operator notices these issues. Deep learning models process several types of telecom data to predict which data are anomalies.

- Call detail records (CDR) data from customer complaints determines if customers have recurrent cellular issues.
- Key performance indicators (KPI) data determines which base stations are down at a given time.
- Twitter tweets identify customers’ dissatisfaction on social media that isn’t directly reported to the company.

Technology & Workflow

Keras: Open-source neural network library. The deep learning engine to train the model.

Google AI: Pre-training language model. The embedding layer.

VaderSentiment: Sentiment analysis algorithm for social media content. The sentiment analysis tool.


Grafana: Open source metric analytics & visualization suite. Visualize the anomaly results and twitter sentiment analysis results.

Anomaly Detection - KPI

<table>
<thead>
<tr>
<th>Input Shape</th>
<th>Dimension Reduction</th>
<th>LSTM</th>
<th>AutoEncoder</th>
</tr>
</thead>
<tbody>
<tr>
<td>39 Dimensions</td>
<td>200</td>
<td>200</td>
<td>36</td>
</tr>
<tr>
<td>Output Shape</td>
<td>39 Dimensions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparing the Upload Traffic Volume, we plotted the input data in blue vs. the prediction data from our model in orange.

We calculated the reconstruction error between the two datasets and determined the anomaly threshold for this feature to identify the anomalies.

Anomaly Detection - CDR

<table>
<thead>
<tr>
<th>Input Layer</th>
<th>Hidden Layer</th>
<th>Output Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>KPI-Like Data</td>
<td>23G data</td>
<td>2/3G Anomaly Result</td>
</tr>
<tr>
<td>KPI-Like Data</td>
<td>4G VoLTE data</td>
<td>4G Anomaly Result</td>
</tr>
</tbody>
</table>

Twitter Sentiment Analysis

Twitter API

Customer Issue Content

Live streaming data

Bert Embedding

Input Shape

LSTM Layer

Dense Layer (256 units, relu)

Dense Layer (128 units, relu)

Classification Result

Tweet with sentiment value

VaderSentiment

Comparison of UPLOAD TRAFFIC VOLUME

Comparing the Upload Traffic Volume, we plotted the input data in blue vs. the prediction data from our model in orange.

We calculated the reconstruction error between the two datasets and determined the anomaly threshold for this feature to identify the anomalies.