**Journal Report 14**

**12/1/23**

I spend my time researching VAEs and their weaknesses in relation to time series anomaly detection. I read these article: <https://www.geeksforgeeks.org/auto-encoders/> and <https://www.dremio.com/wiki/variational-autoencoders/#:~:text=Representation%20Learning%3A%20VAEs%20learn%20a,clustering%2C%20classification%2C%20or%20visualization>.

These are my notes after reading these articles.

Strengths of VAE in Time Series Anomaly Detection

* Effective Representation Learning: VAEs efficiently learn a compressed representation of the data, capturing the essential features and filtering out noise. This representation can be used to identify anomalies that deviate significantly from the normal patterns.
* High Sensitivity: VAEs are sensitive to subtle changes in the data, making them effective at detecting even small anomalies that might be missed by other methods.
* Flexibility: VAEs can be easily adapted to different types of time series data by adjusting the network architecture and training parameters.
* Interpretability: The latent variables learned by VAEs can be interpreted to understand the underlying structure of the data and gain insights into the causes of anomalies.
* Generative Capability: VAEs can be used to generate new data points that resemble the normal patterns, which can be helpful for anomaly detection by comparing the real data to the generated data.

Weaknesses of VAE in Time Series Anomaly Detection

* High Computational Cost: Training VAEs can be computationally expensive, especially for large datasets or complex models.
* Sensitivity to Outliers: VAEs can be sensitive to outliers in the training data, which can lead to inaccurate anomaly detection results.
* Black Box Nature: While the latent variables learned by VAEs can be interpreted to some extent, the internal workings of the model can be difficult to understand.
* Difficulty Capturing Long-Term Dependencies: Standard VAEs may struggle to capture long-term dependencies in time series data, which can limit their ability to detect certain types of anomalies.
* Limited Detection of Contextual Anomalies: VAEs primarily focus on detecting point-level anomalies, and they may not be effective at detecting anomalies that are context-dependent or require knowledge of the overall sequence.

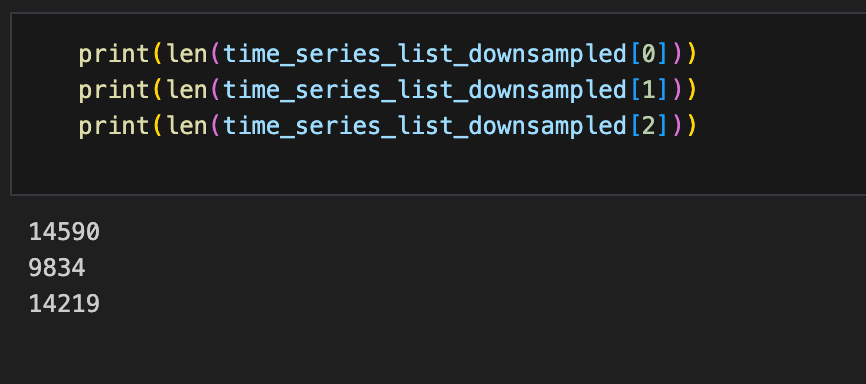
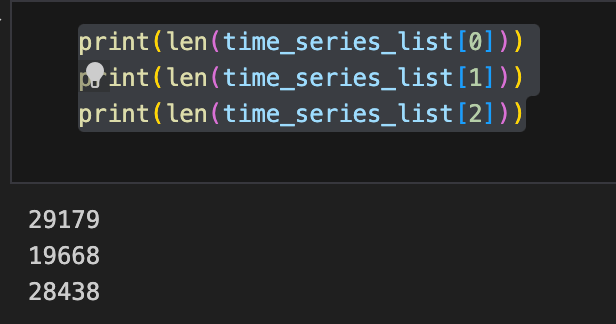
In summary, VAEs cannot accurately create synthetic data for my specific purposes because I want to preserve the contextual anomalies within my dataset and be able to detect them. Furthermore, I am focused on low computational cost algorithms because the end goal is to enable anomaly detection forecasts and predict anomalies with relatively low latency.

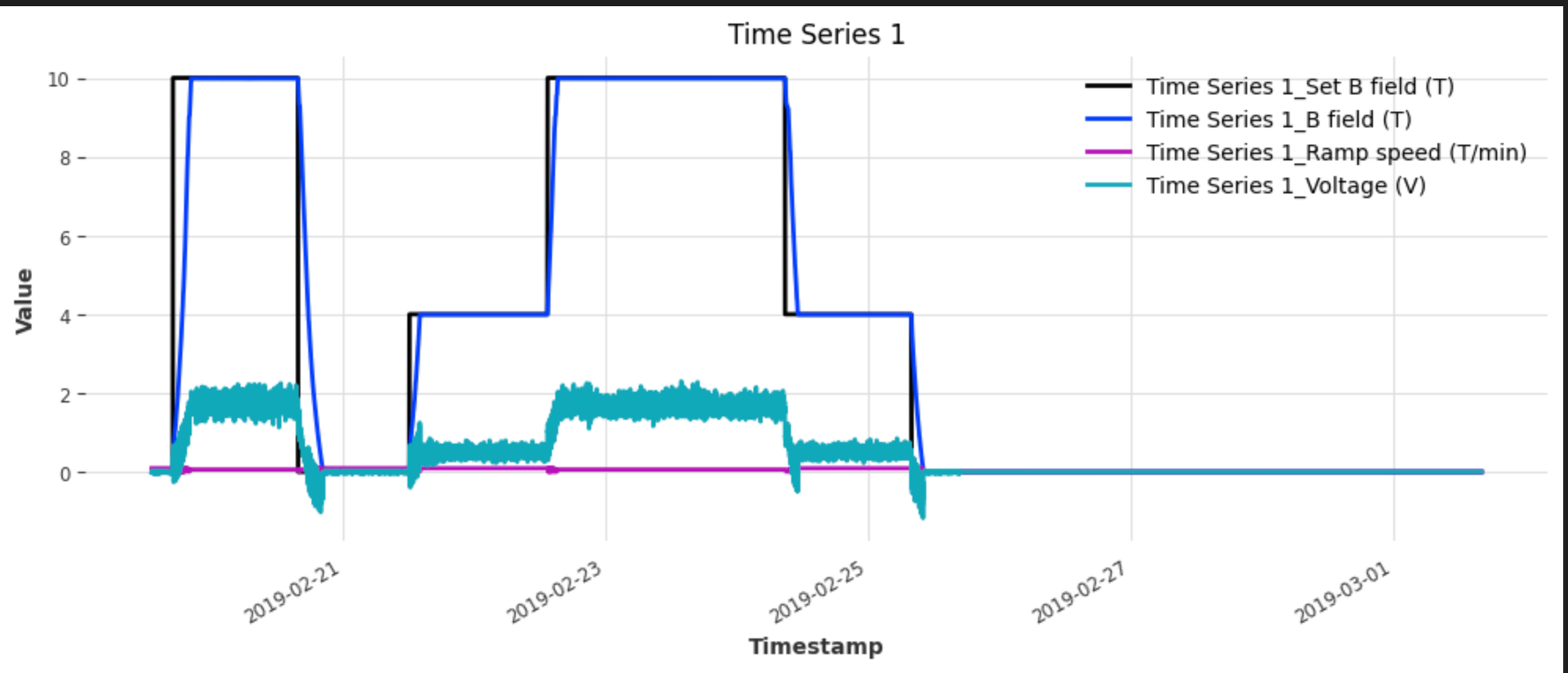
**12/4/23**

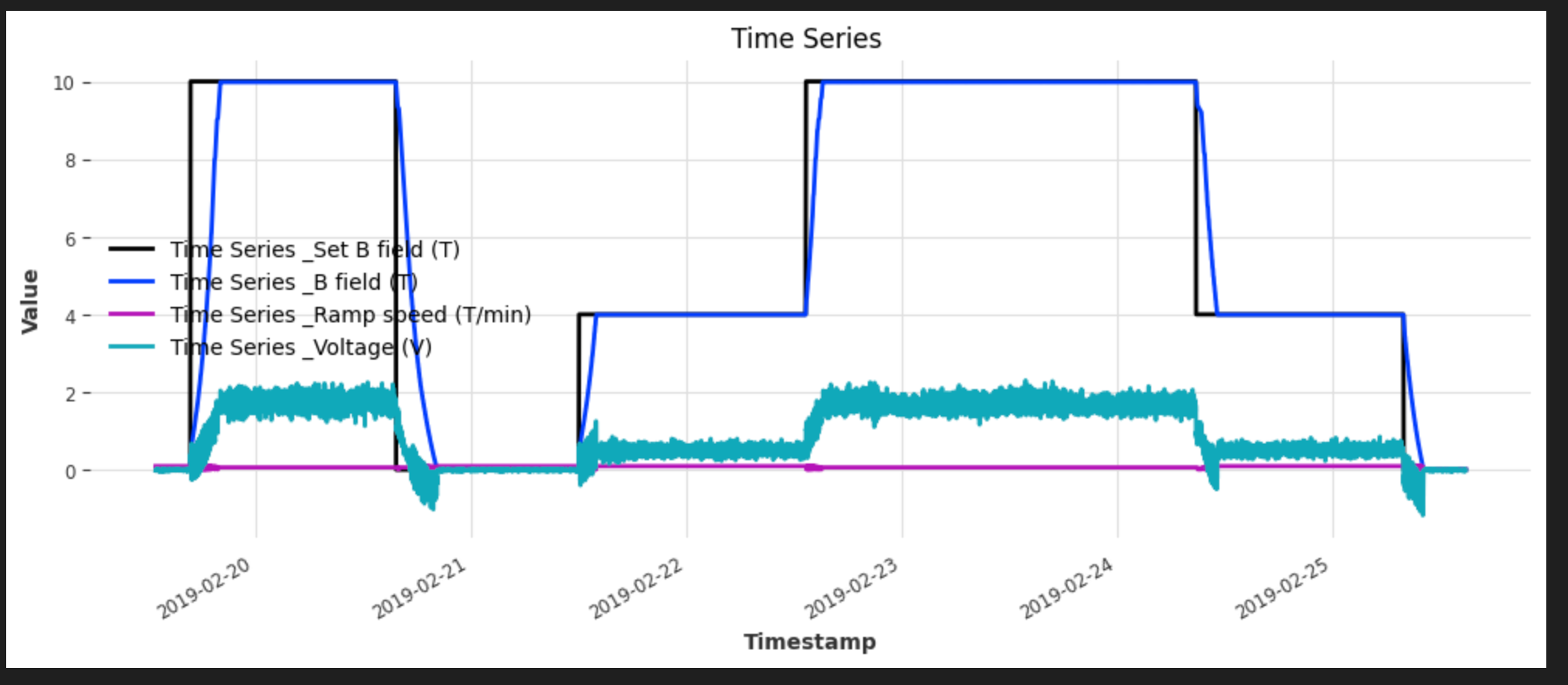
I spent the class time looking into various unsupervised anomaly detection models. I spent about 20 minutes looking, but didn’t find models that fit my task. I am planning to continue looking over the next couple of classes. Afterwards, I loaded my dataset into VSCode and prepared a jupyter notebook to do some preprocessing for my data.

**12/6/23**

I spent the class time preparing my data to be fed into a transformer model. However, I want to minimize the training time and create an initial model for pure anomaly detection. I selected the 4 features that I had originally intended to train my data synthetic model on. I executed two additional steps, which was taking out long instances of constant values and resampling my data again into 1 minute intervals. The picture on the left shows the lengths of the time series before downsampling and the right after downsampling. I am researching methods of taking out the segments in the beginning and end that show a constant value as shown by the graph below. Furthermore, I hope to implement some kind of mapping in the future that would allow me to get the high dimensionality of my dataset into a low dimensionality. I found Self Organizing Maps (SOMs), which looks promising.



****

****