**Using Machine Learning for Anomaly Detection**

Ryan W. Park

Thomas Jefferson High School for Science and Technology

Computer Systems Laboratory

Dr. Peter Gabor

May 29, 2024

**Abstract**

This research addresses the challenge of manual anomaly detection within the National Institute Standards for Technology (NIST) Center for Neutron Research, focusing on the gas compression-based closed-cycle refrigerator (CCR). Currently, temperature control issues are identified through manual troubleshooting. It is crucial to transition from manual to automated data analysis systems due to the constraints surrounding neutron beam experiments. Researchers must reserve beamtime, the allocated time for these facilities, years in advance. These experiments typically last about one week, leaving little tolerance for errors or interruptions. CCR malfunctions in the current manual processes may limit the effective utilization of the scarce beamtime available, resulting in an inefficient use of taxpayer funding for these resources. This study proposed an automated data analysis system that utilizes a neural network to automatically learn patterns in equipment monitors and detect unusual measurements. Specifically, this study aims to implement a transformer model to construct behavior models for control parameters and sensor channels, facilitating continuous monitoring and comparison of predicted sequences with observed data. The goal is to establish comprehensive tracking and prediction capabilities for all aspects of the experiment, consolidated within an integrated beamline status dashboard and text-alert system. This research creates an anomaly detection pipeline for CCR experiments that is adaptable to other sensor channels using unsupervised machine learning. Additionally, this study encompasses the massaging of data into a standardized format suitable for integration with diverse anomaly detection systems. Future work will integrate additional signal and control monitoring so that users of the NCNR can swiftly detect errors, enabling them to take immediate action and significantly reduce the unnecessary consumption of valuable beamtime.

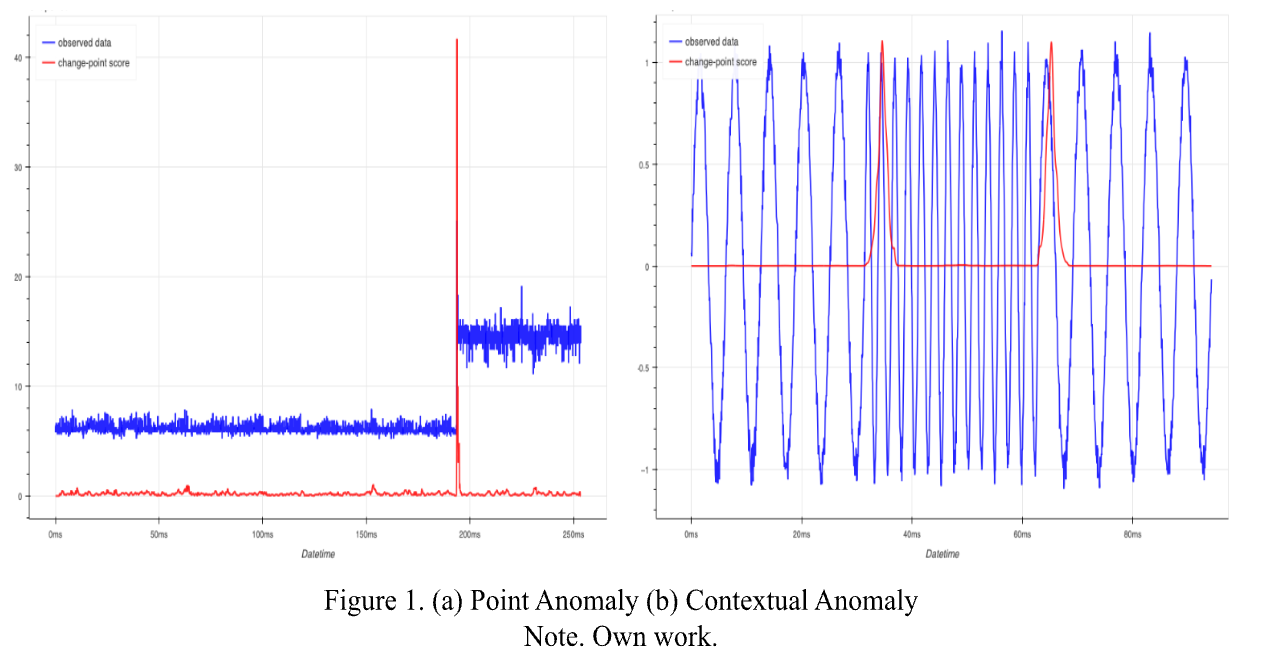
*Keywords:* anomaly detection, transformers, data aggregation, closed-cycle refrigerator

**Introduction**

Continuous operation of real-world systems generates copious sequential measurements recorded across multiple sensors, including those embedded in industrial machinery and space exploration equipment. Detecting anomalies within such extensive system monitoring data is crucial for enhancing security and mitigating financial losses. However, anomalies are typically infrequent and camouflaged amidst the vast expanse of normal observations, rendering data labeling an arduous and inefficient task. Hence, this research focuses on unsupervised learning techniques for time series anomaly detection.

**Figure 1**

*(a) Point anomaly (b) contextual anomaly*



Time series anomaly detection aims to identify abnormal subsequences of varying lengths within these datasets. Thresholding, one of the most elementary techniques, identifies data points breaching predetermined normal ranges. Figure 1 illustrates two distinct anomaly types effectively addressed through thresholding: point anomalies, characterized by abrupt discontinuities at specific timestamps, and contextual anomalies, relying on information from previous time periods. While contextual anomalies are often visually discernible and can be subjected to thresholding using equations (Tuli et al., 2022), a portion of anomalies do not breach preset boundaries. Instead, they manifest as ostensibly "normal" values but assume an unusual character when considered within their temporal context, giving rise to contextual anomalies posing a more intricate detection challenge due to ambiguous contextual signals (Geiger et al., 2020).

To enhance threshold-based anomaly detection, statistical approaches have been introduced, including Statistical Process Control, which identifies anomalies when data points deviate from statistical hypothesis tests. However, these models require human expertise for setting initial assumptions (Geiger et al., 2020).

Classic unsupervised anomaly detection methods, such as density-estimation techniques (e.g., local outlier factor), clustering-based approaches (e.g., one-class SVM, SVDD) (Xu et al., 2021), have been proposed. However, these methods typically lack temporal information and do not generalize well to unseen scenarios. The Anomaly Transformer represents a recent innovative approach, introducing the concept of association discrepancy as a pivotal observation and resulting in Anomaly Attention, a two-branch structure designed to effectively encapsulate this discrepancy (Wen et al., 2023). Recent empirical studies have sought to enhance the informativeness of time point associations through transformer models, demonstrating the Anomaly Transformer's state-of-the-art results.

Machine learning models, both prediction-based and reconstruction-based, are sensitive to irregular and anomalous instances within the training data, particularly for time series anomaly detection (Ren et al., 2019). To mitigate this issue, researchers employ preprocessing techniques, encompassing data normalization and cleaning. Data normalization, applied to training and testing datasets, ensures uniform scaling of attribute values across all variables based on maximum and minimum values from the training data, facilitating analysis and interpretation (Zhao et al., 2020).

Preprocessing techniques for multivariate data differ from those for univariate data. However, the described normalization method is applicable to both and maintains a standardized approach. Additionally, the TadGAN algorithm, a Time Series Anomaly Detection technique employing Generative Adversarial Networks (GANs), utilizes a detrending function to reduce noise and smooth data (Geiger et al., 2020). This detrending function involves subtracting results of a linear least squares fit to the signal, particularly valuable for time series datasets exhibiting linear trends (Geiger et al., 2020). Furthermore, TadGAN incorporates Principal Component Analysis (PCA), a dimensionality-reduction technique for data reconstruction, limited to linear reconstruction and requiring data to exhibit high correlation and follow a Gaussian distribution (Geiger et al., 2020).

The National Institute of Standards and Technology (NIST) and the National Center for Neutron Research (NCNR) are prominent pillars in American scientific research and technological advancement (NIST, 2017). Under the U.S. Department of Commerce, NIST plays a pivotal role in developing and disseminating precision measurement standards, innovative technology, and vital support to industrial sectors. Within NCNR, the Sample Environment Group exemplifies a crucial research facet, managing environmental conditions during neutron beam experiments. Depicted by Figure 2, an apparatus under their purview is the Closed Cycle Refrigerator (CCR), a device capable of cooling samples to -452 degrees Fahrenheit. Driven by the Non-Equilibrium Initiative's vision, NCNR optimizes operational processes, reducing inefficiencies and streamlining endeavors. This pursuit underscores commitment to excellence and efficiency within neutron research, exemplifying this introductory exposition.

**Figure 2**

*Front and back of Closed Cycle Refrigerator CNA-31 at the NCNR*

**A machine with a fan on it

Description automatically generated**A machine in a factory

Description automatically generated

*Note. The image was taken from TOP LOADING CCRs | NIST. (2017, June 6). NIST. https://www.nist.gov/ncnr/sample-environment/equipment/closed-cycle-refrigerators-ccr/top-loading-ccrs*

**Methods**

**Data**

The dataset underpinning this research endeavor was obtained from the Sample Environment Group of the National Center for Neutron Research (NCNR). This group plays an integral role in facilitating comprehensive support for researchers across various stages of neutron beam experiments, encompassing experimental planning, sample loading, equipment preparation, and spectrometer mounting. Over the past three decades, the Sample Environment Group has meticulously archived data from neutron beam experiments, encompassing a diverse array of information, including data pertaining to the Closed Cycle Refrigerator (CCR).

However, it is imperative to note that the raw data acquired from these experiments lacked normalization and standardization. To address this, a Python script was created to extract, organize, and refine the dataset attributes, ultimately culminating in the generation of a pickled .csv file named "CCR.csv." This dataset encompasses records from 45 unique experiments conducted over a 20-year period, and it is noteworthy that this dataset was compiled without any prior screening for anomalies.

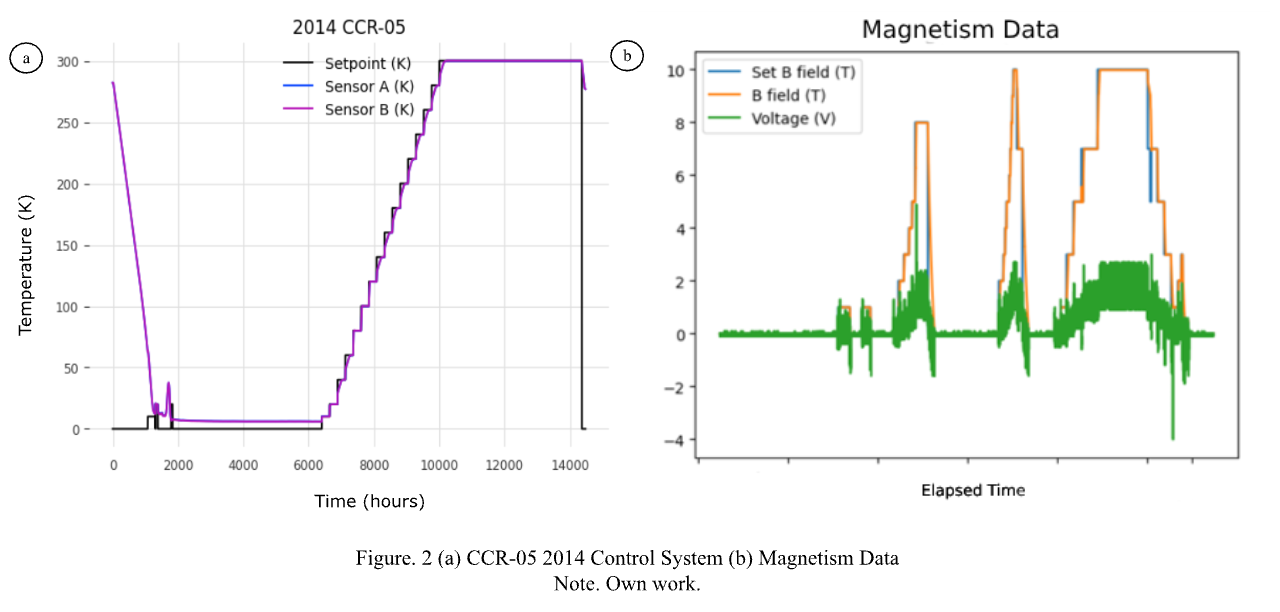
To facilitate the differentiation of time series originating from distinct experiments within the dataset, an "ID" attribute was introduced. This "ID" attribute assigns a numerical value to each instance of a time series, aligning it with the corresponding experiment. For instance, all instances within the time series of the initial experiment bear the value "1" in the "ID" attribute. The diverse attributes featured within this dataset include Elapsed Time (minutes), Setpoint (K), Heater Output % (0-100), Slope, Intercept, Sensor A (K), Sensor B (K), Sensor C (K), and Sensor D (K).

Considering the limited number of experiments available within the CCR dataset, a decision was made to assemble an ancillary dataset that exclusively centers on magnetism-related data. This decision was predicated upon the observed alignment between the magnetism dataset and the CCR dataset, both in terms of prevalent trends and common attributes. This concordance is visually represented in Figure 3, where one can observe analogous attributes, such as "Set B Field (T)" and "B Field (T)," exhibiting resemblance to "Set Point" and "Sensor A," respectively.

This supplemental dataset was procured from the National Center for Neutron Research (NCNR) via a procedure analogous to the acquisition method employed for the primary dataset. However, it is imperative to underscore that this dataset specifically encompasses data emanating from a double-focusing thermal triple-axis spectrometer. The resultant dataset, designated as "magnetism.csv," encompasses an extensive compendium of data gleaned from 2000 experiments conducted over a chronological expanse of 23 years. Notably, the magnetism dataset bears a distinctive hallmark in that each experiment is intricately linked with a unique control sequence, thereby demarcating it from the CCR dataset. The attributes featured within the magnetism dataset encompass Timestamp, Elapsed Time (minutes), Set B Field (T), B Field (T), Inner Lower Temperature, Outer Upper Temperature, Voltage (V), Outer Lower Temperature, 1st Stage Temperature, Shield Temperature, Ramp Speed (T/min), Inner Upper Temperature, and 2nd Stage Temperature.

**Figure 3**

1. *CCR-06 2014 Control System (b) Magnetism Data*

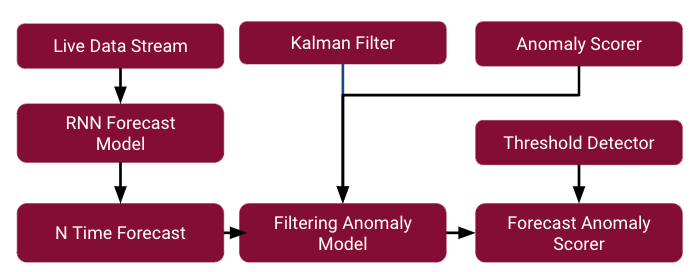
****

**Pipeline**

This study implements a multi-model pipeline for time series anomaly detection in the context of the Closed Cycle Refrigerator (CCR) experiments conducted by the National Center for Neutron Research (NCNR). The pipeline consists of three main components: a recurrent neural network (RNN) for forecasting expected CCR behavior, a filtering anomaly model, and an anomaly scoring model for identifying divergences from expected patterns.

**Figure 4**

*System Summary for Anomaly Detection*



As shown in Figure 4, the pipeline commences with the direct input from the CCR, which then undergoes processing via the RNN Forecast model. This model generates a time forecast covering the remaining duration of the experiment, denoted as "N" timestamps. Concurrently, the Kalman Filter and Anomaly Scorer modules are initialized. The Anomaly Scorer comprises the K-Means Scorer and Wasserstein Scorer components. The Filtering Anomaly Model is then trained using these inputs to discern normal behavior. Finally, in conjunction with the Threshold Detector, the Forecast Anomaly Scorer is tasked with pinpointing and labeling the precise timestamps identified as anomalous data.

A pivotal objective was the identification of the target variable to be forecasted. Consequently, a forecasting model was developed with the aim of capturing the intrinsic relationship between the "Set B Field" and characteristics of interest. An RNN architecture was employed, effectively learning the expected pattern where the target "B Field" closely follows the user defined "Set B Field" values. The process involved systematically evaluating different combinations of past covariates and targets to determine the optimal configuration. The selection criteria prioritized model performance in accurately forecasting the dynamics of interest. The final configuration leveraged the voltage and ramp speed variables as past covariates, with the setpoint as the sole future covariate input. A brute-force approach was adopted to thoroughly assess potential combinations, ensuring the model's ability to robustly forecast the salient characteristics of the system's behavior. This model was initiated with these hyperparameters:

input\_chunk\_length=400 and output\_chunk\_length=100.

The forecasted values generated by the RNN model were subsequently inputted into the Filtering Anomaly Model, the core component responsible for anomaly identification. This model integrates three key elements: the forecast from the RNN, a Kalman filter for noise reduction, and an Anomaly Scorer for quantifying deviations from expected patterns. The Kalman filter component employs a recursive algorithm to interpret the time series data as observations from a potentially noisy linear dynamical system with an underlying hidden state. By recursively updating its state estimate based on the incoming observations, the Kalman filter effectively reduces the noise levels present in the forecasted data, yielding a denoised signal that more accurately represents the underlying system dynamics. The Anomaly Scorer plays a pivotal role in quantifying the divergence between the denoised forecast and the observed data at each time step. It achieves this through the application of mathematical scoring functions that measure the dissimilarity between the two time series representations. By computing these anomaly scores, the model can identify time points where the observed data deviates significantly from the predicted normal behavior, indicating the presence of potential anomalies.

The KMeans Scorer plays a crucial role in identifying local anomalies within the time series data. This technique computes the distance between each data vector of size W and the nearest centroid from a set of k predetermined centroids. By quantifying the divergence from these local distribution centers, the KMeans Scorer effectively captures anomalies that manifest as deviations from typical local patterns observed during the training phase. This approach is particularly well-suited for analyzing localized distributional shifts that may be indicative of transient anomalous behavior.

In contrast, the Wasserstein Scorer is designed to detect global distributional anomalies that transcend local fluctuations. It achieves this by calculating the Wasserstein distance between the training distribution, which represents the expected normal behavior, and the observed distribution at each time step. This distance metric quantifies the minimum cost of transforming one distribution into the other, making it highly sensitive to significant shifts in the overall data distribution. The Wasserstein Scorer's ability to capture these global deviations complements the local sensitivity of the KMeans Scorer, enabling a comprehensive evaluation of potential anomalies.

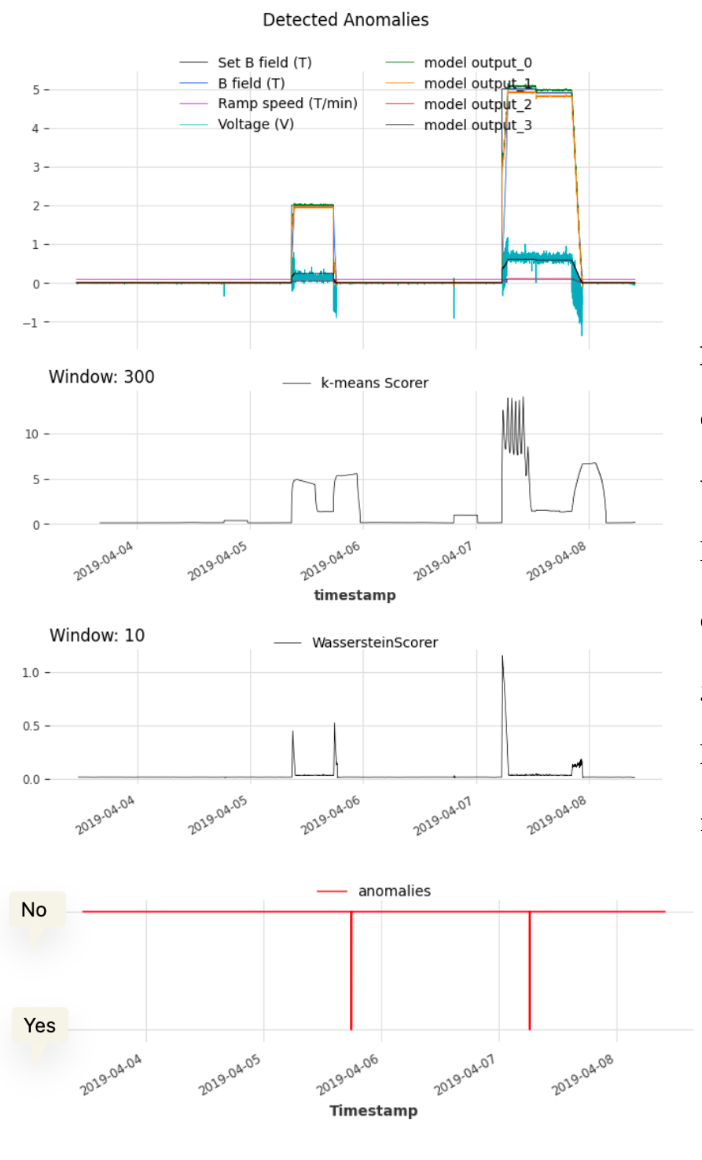
The Filtering Anomaly Model serves as the central component, integrating the outputs from the KMeans and Wasserstein Scorers to identify and remove irrelevant data points. Underpinning this model is the assumption that it can effectively filter the time series when no anomalies are present. To ensure satisfactory performance, the Filtering Anomaly Model is trained exclusively on anomaly-free time series, establishing a robust baseline for normal behavior. During the inference phase, the model leverages the anomaly scores from both the local and global scorers to systematically filter out data points that deviate significantly from the expected patterns, while preserving the relevant portions of the time series for further analysis.

The anomaly scores computed by the Anomaly Scorer are further processed by a threshold detector, which serves as the final stage in the anomaly identification pipeline. This component applies a predefined threshold to the anomaly scores, effectively labeling any time points with scores exceeding the designated threshold as anomalies. The output of the threshold detector is a new time series, where each timestamp is marked with a binary label indicating whether it represents an anomaly ("yes") or not ("no"). Figure 5 provides a visual representation of this output, with anomalous time points clearly distinguished from the normal observations.

By incorporating the threshold detector, the methodology establishes a systematic approach to translating the continuous anomaly scores into discrete anomaly labels. This step is crucial for practical applications, as it enables the clear delineation of anomalous regions within the time series data. The specific threshold value can be determined through empirical analysis or domain-specific considerations, allowing for the adjustment of the sensitivity of the anomaly detection process to suit the requirements of the application at hand.

**Figure 5**

*Final Results of the Anomaly Detection Pipeline*



**Results**

**Variable Forecasting and Anomaly Detection**

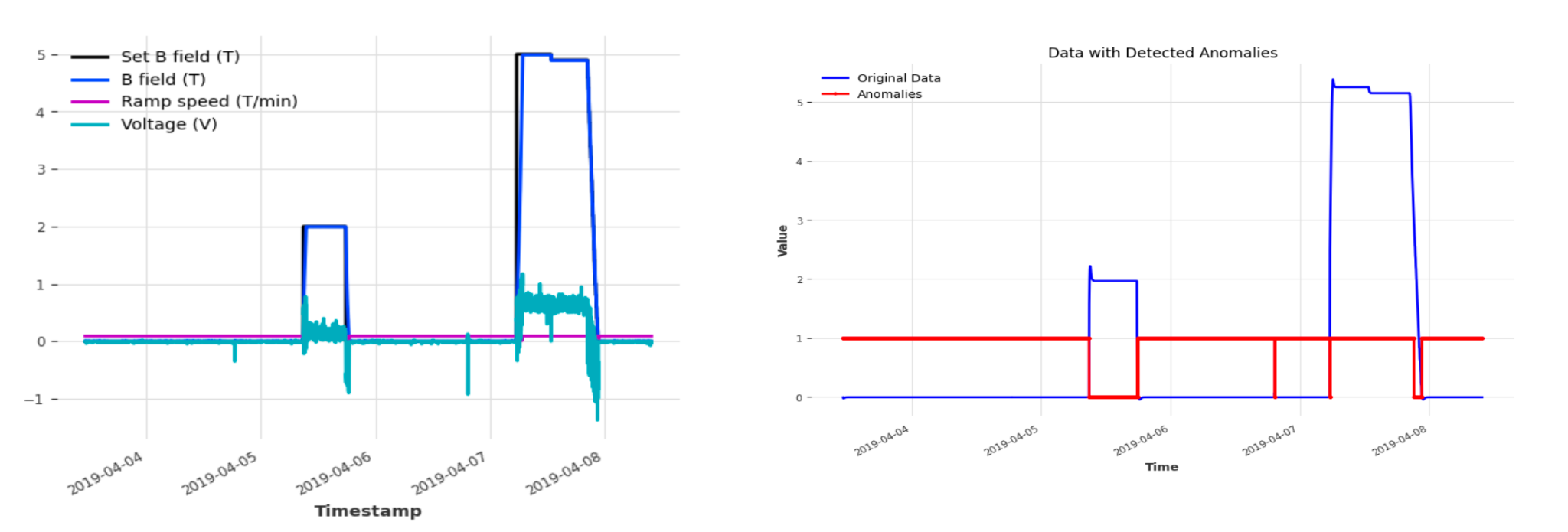
Figure 6 presents the conclusive outcomes of the proposed pipeline. The initial graph in this figure depicts the B-field, representing the forecasted variable using the Recurrent Neural Network (RNN) Forecasting Model. This output reflects the anticipated behavior generated by the model. The middle graph shows the K-Means score at each timestamp, highlighting its high sensitivity to abrupt data changes. This sensitivity is expected due to the changes in the Controlled Current Regulator's (CCR) set point. The lower graph illustrates the precise detection of anomalies at distinct timestamps.

**Preliminary Model Testing**

Figure 6 demonstrates the preliminary testing of the model, which initially employed only a threshold detector. Anomalies were identified when values exceeded or fell below a specified threshold. This approach revealed a significant weakness: it was overly sensitive to large data changes and did not account for set point variations, leading to misclassification of reasonable sensor value fluctuations as anomalies. The threshold detector failed to learn from previous data patterns, which is crucial for the CCR where long-term anomalies are of interest and set point changes are frequent due to varying experimental requirements.

**Figure 6**

*Initial Testing with Only Threshold Detector and No Scorer*

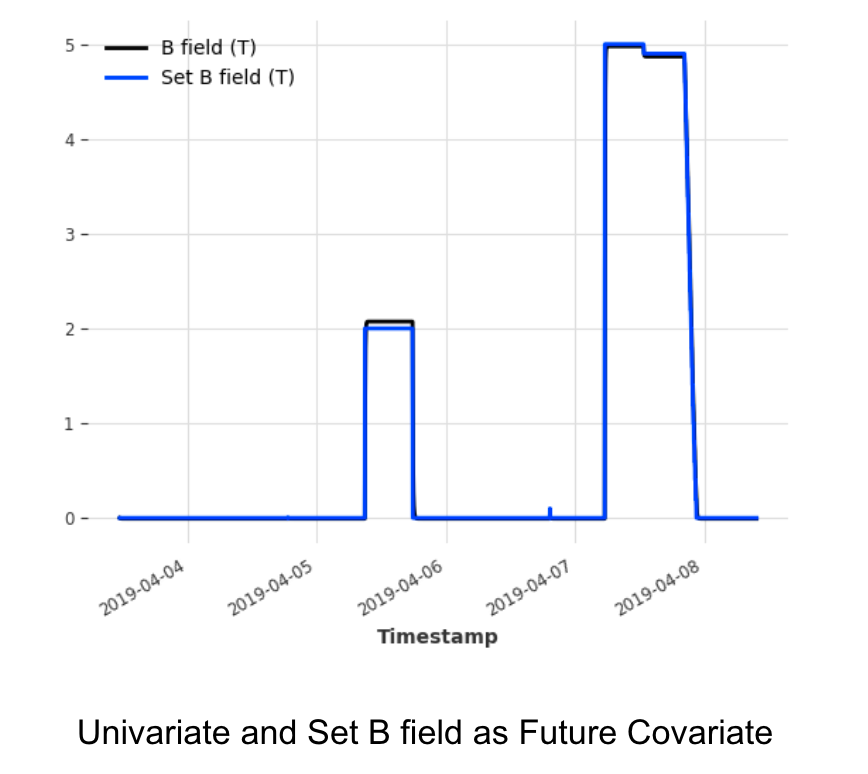
****

**RNN's Pattern Recognition Capability**

Figure 7 highlights the RNN's capability to learn and recognize the patterns in the CCR data stream. The model successfully identified that the B-field should closely follow the set B-field but should not directly overlap, exhibiting slight fluctuations instead. This understanding is essential for accurately modeling the CCR's behavior.

**Figure 7**

*Univariate and Set B field as Future Covariate*

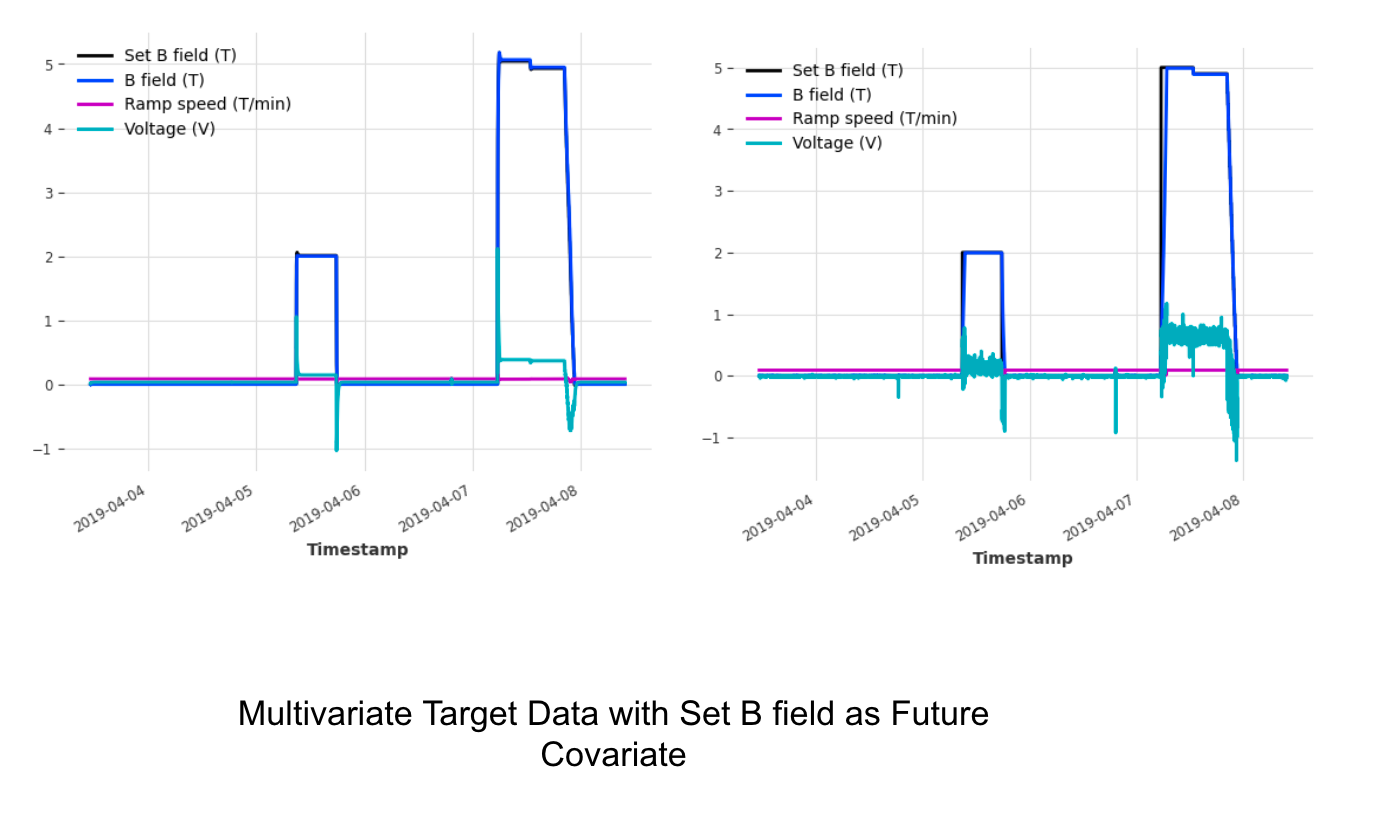
****

**Finalized RNN Forecast Model**

Figure 8 showcases the finalized RNN forecast model, which effectively captures the noise and non-linearity inherent in the CCR's parameters. This comprehensive modeling is logical since all other variables are dependent on the set B-field, and the RNN is robust enough to understand these complex relationships.

**Figure 8**

*Multivariate Target Data with Set B field as Future Covariate*

****

**Discussion**

A novel anomaly detection pipeline has been developed to autonomously identify and notify users of anomalies within the Cold Source Refrigerator (CCR) during neutron beam experiments. This pipeline demonstrates the capability to detect various types of anomalies, including those requiring analysis of past patterns, thereby addressing the project objectives. While the project has yielded promising results, there are several areas for potential improvement. The lack of labeled datasets has posed challenges in calculating comprehensive model evaluation metrics. Additionally, logistical constraints have limited thorough testing with live CCR data, as accessing the CCR located in Bethesda, Maryland, for testing purposes is impractical. Furthermore, there were logistical constraints with the hyperparameter tuning of the Forecast RNN Model and Filtering Anomaly Model. The lack of computational resources prevented rigorous testing as training time was constrained. Training for each iteration of the model exceeded 6 hours and I felt that it was better time allotted for the training of the Filtering Anomaly Model. Furthermore, the implementation of more anomaly scorers is necessary to capture more anomalous patterns within the data stream. Specifically, the Wasserstein scorer and K-means scorer are efficient for global and local minimum and maximums but are insufficient for capturing long-term anomalies prevalent in CCR data. This pipeline needs to be further tested with live data from CCR at NCNR to ensure the proper execution and upholding of computational prediction time during realistic use scenarios. I did not have access to the CCR as this was a remote project from the NCNR during the school year.

In terms of future work, there is room for more focus on directly integrating the anomaly detection pipeline with the CCR hardware and implementing a notification system to alert users of anomalies through an app or other means. It is imperative that there is a UI for notification of these anomalies as the NCNR seeks to adopt an automatic data analysis system to pinpoint the exact timestamp in which an anomaly occurs. Although challenges remain, the development of an effective anomaly detection pipeline represents a significant advancement. Continued efforts to enhance integration with the hardware and address logistical constraints could further improve the performance and applicability of the anomaly detection system.

**References**

Closed Cycle Refrigerators (CCR). (2017). *NIST*.<https://www.nist.gov/ncnr/sample-environment/equipment/closed-cycle-refrigerators-ccr>

Geiger, A., Liu, D., Alnegheimish, S., Cuesta-Infante, A., & Veeramachaneni, K. (2020, September 16). *TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks*. ArXiv.Org.<https://arxiv.org/abs/2009.07769v3>

Ren, H., Xu, B., Wang, Y., Yi, C., Huang, C., Kou, X., Xing, T., Yang, M., Tong, J., & Zhang, Q. (2019). Time-Series Anomaly Detection Service at Microsoft. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3009–3017.<https://doi.org/10.1145/3292500.3330680>

Tuli, S., Casale, G., & Jennings, N. R. (2022). *TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data* (arXiv:2201.07284). arXiv.<https://doi.org/10.48550/arXiv.2201.07284>

Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2023). *Transformers in Time Series: A Survey* (arXiv:2202.07125). arXiv.<http://arxiv.org/abs/2202.07125>

Wu, R., & Keogh, E. (2021). Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress. *IEEE Transactions on Knowledge and Data Engineering*, 1–1.<https://doi.org/10.1109/TKDE.2021.3112126>

Xu, J., Wu, H., Wang, J., & Long, M. (2021, October 6). *Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy*. International Conference on Learning Representations.<https://openreview.net/forum?id=LzQQ89U1qm_>

Zhao, H., Wang, Y., Duan, J., Huang, C., Cao, D., Tong, Y., Xu, B., Bai, J., Tong, J., & Zhang, Q. (2020). *Multivariate Time-series Anomaly Detection via Graph Attention Network* (arXiv:2009.02040). arXiv.<http://arxiv.org/abs/2009.02040>