Existing Models on GitHub

1. Anomaly Detection Pipeline with Kedro : https://github.com/kennethleungty/Anomaly-Detection-Pipeline-Kedro?utm_source=chatgpt.com
2. Log Analysis for Anomaly Detection : https://github.com/davide97l/Log-Analysis-for-Anomaly-Detection?utm_source=chatgpt.com
3. Orion: Unsupervised Time Series Anomaly Detection : https://github.com/sintel-dev/Orion?utm_source=chatgpt.com
4. Anomaly Detection Pipeline on Azure Databricks : https://github.com/devlace/azure-databricks-anomaly?utm_source=chatgpt.com
5. **TravisTorrent Dataset**

- A comprehensive dataset combining Travis CI and GitHub data, encompassing over 2.6 million builds from more than 1,000 projects.

- Includes metadata, build logs, and commit information, making it suitable for training ML models to predict build failures and anomalies.

  https://gousios.org/bibliography/BGZ17a.html?utm_source=chatgpt.com

🤖 Machine Learning Models

1. **Continuous Defect Prediction (CDP)**

   - Utilizes the TravisTorrent dataset to predict risky software changes that could break the build if committed to a repository with CI enabled.

   - Employs various software process metrics as features for defect prediction.

     https://arxiv.org/abs/1703.04142?utm_source=chatgpt.com

2. **BugSwarm**

   - A toolkit that mines and archives failing and subsequent passing runs from Travis CI, creating a scalable dataset of reproducible failures and fixes.

- Facilitates the evaluation of fault-detection, localization, and repair methods.

  https://arxiv.org/abs/1903.06725?utm_source=chatgpt.com

.