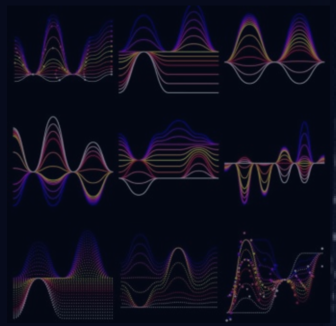
AI4ASD: AI for Autism Spectrum Disorder

Artificial intelligence applied to ASD at Hospital Universitari Mutua Terrassa



Autism Spectrum Disorder

Why is it important to develop this project?

Currently, assessment instruments and treatment design, based on diagnostic ASD symptoms, are evaluated subjectively, relying on information provided by parents or on non-measurable criteria. There is a lack of objective instruments to measure treatment effectiveness and response.

For this reason, it is essential to identify biomarkers in children with ASD, with multiple observability objectives:

- To understand part of the child's clinical spectrum based on the results provided by biomarkers.
- To continuously know the child's state within each context, enabling the development of therapies, monitoring their progression over time, and making comparisons between results with objective, direct measurement in real time.
- To prevent and communicate critical states in the child that are invisible to human senses, such as the detection of behavioral crises, aggressiveness, anxiety, etc.
- To obtain a measurement of the child's spectrum after the application of medication, dosages, and therapies, in order to objectively evaluate the outcomes.

Why was the Mutua Terrassa center selected?

Tertiary ASD inpatient unit at Hospital Universitari Mutua Terrassa:

- First center in Spain focused exclusively on cases of ASD.
- First center allowing family admission.
- First center in Spain with an intensive and comprehensive therapeuti approach, addressing areas that interfere with adaptive functioning.
- First center with specialized staff in autism, providing 24-hour coverage.

- Image monitoring through intelligent cameras with 24-hour motion detection and escalation management, among other features.Center connected with other leading centers and cutting-edge research lines through its director, Amaia Hervas.Contracted center. 95 percent of admissions are funded by Cat Salut.

NTT DATA

Objetivos del proyecto.

OBJECTIVE 1 (Phase 1):
Measure biomarkers in patients at the ASD Center using a smartwatch. Define KPIs based on biomarkers in order to analyze children's states, including improvements, crises, conditions, trends, etc.

- Blood Volume Pulse (BVP)
- Heart Rate (HR)
- Electrodermal Activity (EDA)
- Skin Temperature (TEMP)
- Accelerometer (ACC)
- Inter-Beat Interval (IBI)
- Sleep
- Sweating and movement

The overall objective is to improve patients' quality of life through better identification of their autism profile and its treatment.

OBJECTIVE 2 (Phase 2):
Have access to each child's profile during psychiatric admission, from the clinical area, providing context to the measurements:
Sensory
Emotional
Psychiatric and neurodevelopmental
Psychomotor
Pharmacological
Genetic
This context will be provided by the hospital and reflected in the KPIs by weighting the biomarker measurements themselves. No additional data inputs will be incorporated into the systems that could compromise personal or medical information, due to security and data protection requirements.

OBJECTIVE 3 (Phase 2):
Apply Big Data to obtain statistical data from biomarkers. Apply Artificial Intelligence to detect anomalies in measurements. Provide dashboards and alerts for measurements, including informational escalation levels.

OBJECTIVE 4 (Phase 2):
Analyze whether, throughout the process of the other objectives, any groups of children can be identified based on biomarkers and psychiatric profiles using Artificial Intelligence.

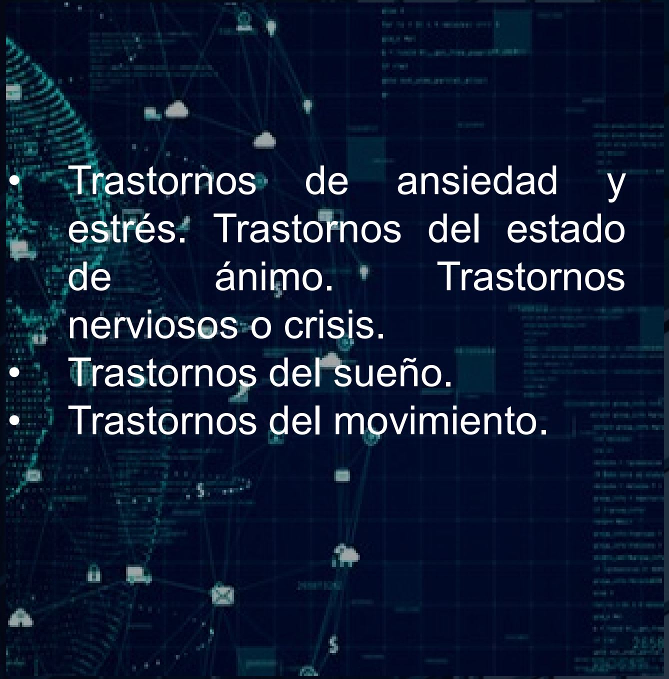
Big Data applied to autism (Phase 1 and 2)

Phase 1: In the first phase, the objective is to continuously measure these biomarkers in the children at the center.

Phase 2: The second objective is to analyze relationships between nursing records and IT system records in order to validate the correlation of events.

DIGITAL BIOMARKERS

- Heart rate
- Heart rate variability
- Respiratory rate
- Usage detection
- Wear detection sensor
- Skin conductance level (EDA SCL)
- Skin temperature
- MET, body position
- Sleep detection
- Step count
- Activity count
- Activity intensity
- Activity classification
- Standard deviation of accelerometer magnitude



- Trastornos de ansiedad y estrés. Trastornos del estado de ánimo. Trastornos nerviosos o crisis.
- Trastornos del sueño.
- Trastornos del movimiento.

WARD-LEVEL RECORDING

- Specialized staff in autism with 24-hour coverage.
- Image monitoring through intelligent cameras with 24-hour motion detection and escalation management, among other features.
- Specialist reports (neurology, psychiatry, therapists) directly linked to ward activity.
- One-to-one monitoring. Individualized reports for children with ASD.

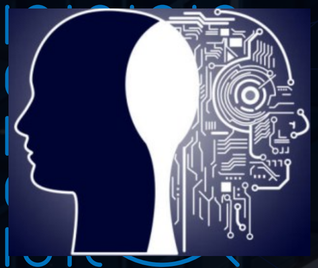
Hypothesis 1: *Digital monitoring will provide metrics that demonstrate the child's states, improving observability.*
Hypothesis 2: *Observability will enable the detection of patterns, anomalies, and trends to be addressed.*

AI applied to autism (Phase 2)

AI will classify all cases based on continuous medical learning. As a result, diagnoses, treatments, medications, and dosages aligned with the autism profile can be preassigned with high precision and at an early stage.

DIGITAL BIOMARKERS

- Heart rate
- Heart rate variability
- Respiratory rate
- Usage detection
- Wear detection sensor
- Skin conductance level (EDA SCL)
- Skin temperature
- MET, body position
- Sleep detection
- Step and activity count
- Activity intensity and classification
- Standard deviation of accelerometer magnitude



***Hypothesis 3:** Artificial intelligence, combined with high-quality field reporting, on-site expert monitoring, and intelligent cameras, can contribute to defining measurement patterns and support nursing reports, 24-hour monitoring, expert assessments, and related activities.*

Neurodevelopmental Profile

- Behavior
- Autism
- Hyperactivity
- Intellectual disability
- Communication
- Learning

Sensory Profile

- High and low thresholds
- Self-regulation
- Active and passive
- Observer
- Seeker
- Sensitive
- Avoidant

Emotional Profile

- Resilience
- Attitude
- Social intuition
- Self-awareness
- Sensitivity to context
- Attention

Clinical Profile

- Blood tests
- Genetics
- Pharmacogenetics and epigenetics
- General psychopathology
- Functionality
- Autonomy

Psychomotor Profile

- Movement
- Rest
- Type of movement
- Intensity
- Stereotypies
- Involuntary movements
- Tics

Project Architecture

Connectivity:
Using a compatible smartphone, the Embrace Plus device sends data to the smartphone via Bluetooth. The data is then sent from the smartphone to Empatica Cloud in 15-minute batches through the smartphone's Internet connection. Raw data (AVRO format) and digital biomarkers (CSV format) are available in Empatica Cloud (AWS S3 bucket). From there, near real-time metrics are generated and subsequently transmitted to the center's on-premise system for specialized AI-based processing.

Security:
Data is encrypted both in transit and at rest. The solution is designed with physical, electronic, and organizational procedures, including encryption, firewalls, access controls, and other safeguards to protect data against loss, misuse, unauthorized access, disclosure, alteration, and destruction, ensuring the protection of user data. Our system complies with HIPAA and GDPR requirements.

IoT Information

Clinical Information

Data access:
Access to the S3 bucket can be carried out in several ways, for example through AWS CLI v2, AWS SDK, or third-party applications such as Cyberduck. A detailed data access guide will be provided, offering further information on the available tools for accessing the AWS S3 bucket. From this point onward, the scope of NTT systems in Azure EU begins.

GDPR
User anonymized data is collected, stored, and processed in order to comply with our contractual obligations with the study investigator. We do not have control or ownership of user data, as it is entirely the user's property. Data would only be accessed for justified reasons, such as resolving an error or issue.

Plan

Start → Phase 1 (3 months) → **3 months** → Phase 2 (6 months) → 8.5 months → End of the project → **Transición (15 dias)** → 9 months

- Smartwatch usage
- Biomarkers implemented
- Demo and deliverable

In case of GO:

- Plan Phase 2
- Review Phase 2 team

- Implementation for 10 admitted patients
- AI-based ASD profile categorization
- System installed, demo, and deliverable

- Handover from NTT to Mutua Terrassa of the project deliverables described in the milestones of each phase

Results

- Technical and scientific feasibility
- Scientific and technological evidence
- Project closure

- Scientific and technological evidence

- End of the project

Milestones

Patient monitoring

- Smartwatch usage
- Biomarkers implemented
- Demo and deliverable
- Collaboration agreement
- Ethics committee approval

AI implementation

- AI patterns based on ASD measurements
- System installed, demo, and deliverable

Handover from NTT to Mutua Terrassa

One-week training

- Final deliverable
- Data deliverable
- Smartphone and smartwatch transferred to hospital ownership
- Azure account delegated to the hospital at the end of Phase 2
- Analysis of next steps for the solution

Scope of the Mutua Terrassa ASD Center

The Mutua Terrassa ASD Center will lead the overall project, be responsible for the scientific component, and define the necessary requirements for NTT to deliver the technological component. It will also act as the information hub, serving as the sole entity responsible for processing medical and personal data.

Mutua Terrassa ASD Center

- Overall project management as the clinical lead of the entire initiative.
- Management of security and compliance regulations with patients.
- Management of IoT equipment usage by the nursing team.
- Definition and interpretation of metrics, conclusions, and proposals for evolution.
- Definition of measurements and scientific validation of conclusions.
- Project approval and authorization for the use of devices and data, as approved by the hospital's ethics committee and clinical studies board.
- Preparation of the medical research project documentation for the procurement of Empatica smartwatches, requiring certification that they are used for medical research purposes.
- Follow-up meetings.
- Project approval and authorization for the use of devices and data, as approved by the hospital's ethics committee and clinical studies board.

Nursing

- Continuous recording of each patient's situations, to be reported to the NTT team.
- Charging, maintenance, and management of smartphones and smartwatches 24x7.
- Monitoring the solution console to perform tests as key users 24x7.
- Follow-up meetings.

Logistics

- Providing access to information systems and related information systems according to project requirements.
- Providing Wi-Fi Internet access to all smartwatches and smartphones.
- Designing a welcome therapy for the smartwatch for children with ASD, to ensure proper use and prevent damage. Managing smartwatches and smartphones to ensure they remain connected and charging.

NTT Scope

NTT will be responsible for developing the technical tasks, under the guidance of the medical team, in order to carry out the measurements.

NTT will carry out the technological tasks required for Phase 1:

- Implement a time-series database that ingests information from smartwatches and represents biomarkers in graphical form.
- The deliverable will be a presentation of the system functionalities to date, along with a document detailing the implementation and the results obtained.
- An estimated maximum duration of 3 months.
- On-site or remote format, as required during the technical tasks of the project.

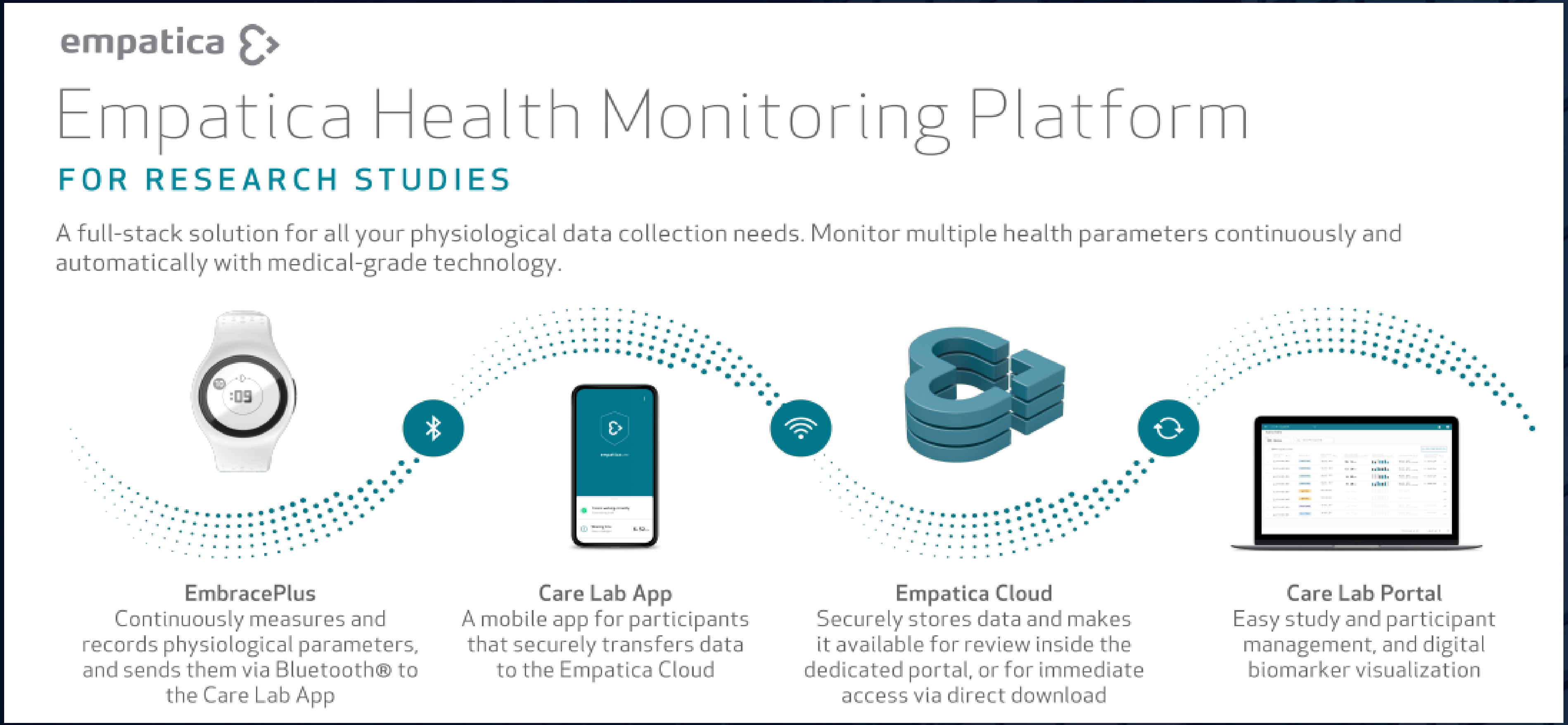
NTT will carry out the technological tasks required for Phase 2:

- Apply Artificial Intelligence to the time-series database for pattern recognition. Generate real-time reports and send highlighted events by email to medical teams.
- The deliverable will be a presentation of the system functionalities to date, along with a document detailing the implementation and the results obtained.
- An estimated maximum duration of 6 months.
- On-site or remote format, as required during the technical tasks of the project.

Scope considerations:

- As a research project, it does not have a defined delivery scope in terms of effectiveness, applicability, or functionality milestones. NTT will be limited to performing the technical tasks of Phases 1 and 2. No additional phases are included.
- No maintenance or warranty is included after the completion of Phases 1 and 2.
- Intellectual property of the NTT FLENS asset is not transferred.
- IoT equipment will be purchased by the hospital and paid for by NTT. A total of 12 devices will be acquired (2 for testing and 10 for the 10 patients at the center).

IoT for Biological Markers



The devices are covered by a 1-year warranty. In the event of any issue, you may contact our research support team, whose objective is to respond within 24 hours. They will conduct an investigation and attempt to resolve the issue. Most problems can be resolved through remote troubleshooting. Otherwise, they will assess the need for a replacement, if applicable. The exact timing of the process will vary from case to case.

IoT for Biological Markers

FDA-CLEARED CE CERTIFIED

embrace PLUS

SENSORS

Empower your research with the world's smallest and most advanced wearable to combine 4 powerful sensors.

PPG
4 CHANNELS (64 HZ SAMPLING RATE)

EDA
0 µSIEMENS - 100 µSIEMENS (4 HZ SAMPLING RATE)

SKIN TEMPERATURE
(1 HZ SAMPLING RATE)

ACCELEROMETER
HIGH PRECISION 3D MICROELECTROMECHANICAL ACCELEROMETER (64 HZ SAMPLING RATE)

FEATURES

EmbracePlus comes with a range of added features that ensure minimal disruption, continuous connectivity, and unparalleled flexibility.

- Bluetooth 5.0
- Compatible with most smartphones (iOS & Android)
- Vibration motor for haptic feedback
- IP67 waterproofing
- Buttons for event tagging
- On-board memory
- Always-on E Ink display

IoT for Biological Markers

Data

DIGITAL BIOMARKERS

Digital biomarkers are 1-minute physiological parameters calculated by Empatica's proprietary algorithms. Available in CSV format.

- Pulse Rate
- Pulse Rate Variability
- Respiratory Rate
- Wearing Detection
- EDA/Skin Conductance Level (SCL)
- Skin Temperature
- METs
- Body Position
- Sleep Detection
- Steps Count
- Activity Counts
- Activity Intensity
- Activity Classification
- Accelerometer Magnitude Standard Deviation

RAW SENSOR DATA

High frequency raw data. Available in Avro format.

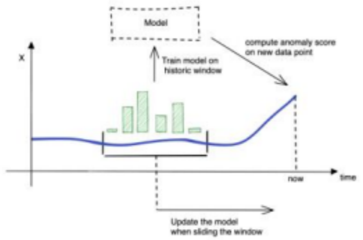
- Accelerometer
- Skin Temperature
- Electrodermal Activity
- Blood Volume Pulse
- Systolic Peaks
- User Tags
- Steps

All data is stored on the GDPR and HIPAA-compliant Empatica Cloud and accessible via an Amazon S3 bucket. **Request sample data.**

NTT DATA

Proposed AI Technologies

Anomaly detection



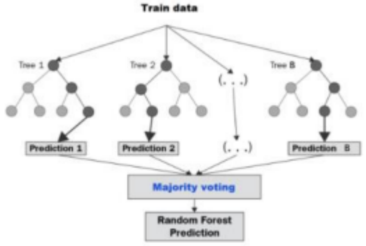
Anomalias en temporalidad meses

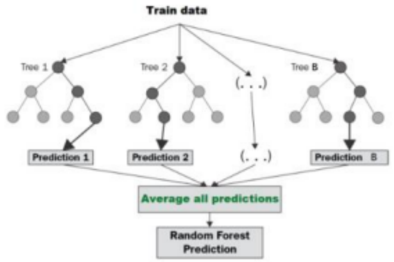
- Uses unsupervised algorithms
- Ensembles traditional-statistical (Local Outlier Probabilities) and deep learning (Long Short Term Memory)
- Combined with Noise Reduction Engine to generate the most meaningful alerts.
- Identifies:
  - Repetitions in a certain period.
  - Time active.
  - Metric correlation with other anomalies.
  - Percentage of metrics higher than a threshold or after smoothihng.

Random forest

- Supervised algorithm for data classification.
- Multiple measurements can be includes to classify data (efficiency increases identifying Principal Components PCA)
- Identifies if data belongs to a group:
  - Regular behavior.
  - Anomalies.
  - Critical.
  - Down.

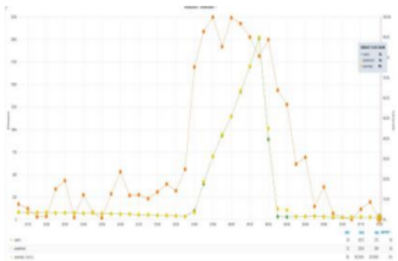
Probabilidad de secuencialidad





Auto-Regressive Integrated Moving Average (ARIMA):

patrones de temporalidad a un año. 3 meses minimos.

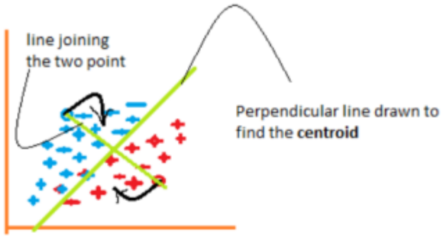


- It's unsupervised algorithm
- Handles data with some seasonality.
- Predicts a measurement behavior and upper and lower confidence interval values:
  - Lower N prediction.
  1. Mean prediction.
  2. Upper N prediction.

K-Means (Clustering):

- Unsupervised algorithm for data clustering.
- Classifies data using clusters.
- Sorts data in groups based in measurement behavior.

Grupos y subgrupos dinamicos



Proposed AI Technologies

AnomalyDetectionCo × AnomalyDetectionED ×

Edit in VS Code · Compute: Serverless Spark Compute - Available · Language: PySpark (Python)

Session not started · Configure session

7.4 Custom Runtime: Overcoming Azure ML Limitations

Azure Machine Learning (AML) does not natively support TensorRT within its default Python environments or Compute Instance. Given that TensorRT offers substantial performance benefits for LSTM inference through kernel fusion and FP16 execution, this limitation would significantly bottleneck the model deployment if left unaddressed.

To overcome this, we developed a **custom Docker image** that packages the required runtime:

- Python 3.9 and Miniconda support
- TensorFlow 2.15 with GPU and TF-TRT integration
- CUDA 12.2
- cuDNN 8
- TensorRT runtime libraries
- Supporting scientific libraries: NumPy, Pandas, etc.

This image is hosted in Azure Container Registry (ACR) and used in AML training jobs and inference pipelines. It allows us to leverage TensorRT even in environments where native support is unavailable.

> The containerized approach ensures portability across compute environments and reproducibility for clinical audits.

Dockerfile Overview

```
# Base image with full CUDA dev tools (including ptxas, libdevice, nvcc)
FROM nvidia/cuda:12.2.2-cudnn8-devel-ubuntu20.04

# Set noninteractive frontend
ENV DEBIAN_FRONTEND=noninteractive \
    TZ=Etc/UTC \
    # Set CUDA location for TensorRT
    LD_LIBRARY_PATH=/usr/lib/x86_64-linux-gnu:$LD_LIBRARY_PATH

# Set up environment variables for XLA (optional, safe defaults)
ENV XLA_FLAGS=--xla_gpu_cuda_data_dir=/usr/local/cuda

# Install Python 3.9, pip, and essential build tools
RUN apt update && apt install -y \
    python3.9 python3.9-dev python3.9-distutils \
    # TF-TRT (TensorFlow integration with TensorRT)
    libnvinfer8 libnvinfer-dev libnvinfer-plugin8 \
    libnvonnxparsers8 libnvparsers8 libnvinfer-bin \
    wget curl git build-essential software-properties-common \
    && rm -rf /var/lib/apt/lists/* \
    && ln -s /usr/bin/python3.9 /usr/bin/python \
    && curl -sS https://bootstrap.pypa.io/get-pip.py | python

# Upgrade pip and install TensorFlow with GPU support + TensorRT support
RUN pip install --upgrade pip && \
    pip install tensorflow==2.15.0 tf2onnx

# Miniconda
RUN wget https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh -O miniconda.sh && \
    bash miniconda.sh -b -p /opt/conda && \
    rm miniconda.sh && \
    ln -s /opt/conda/etc/profile.d/conda.sh /etc/profile.d/conda.sh && \
    echo ". /opt/conda/etc/profile.d/conda.sh" >> ~/.bashrc && \
```

Proposed AI Technologies

Microsoft Foundry | Azure Machine Learning

Azure subscription 1
biogen-nonprod-aml

BioGen > biogen-nonprod-aml > Notebooks

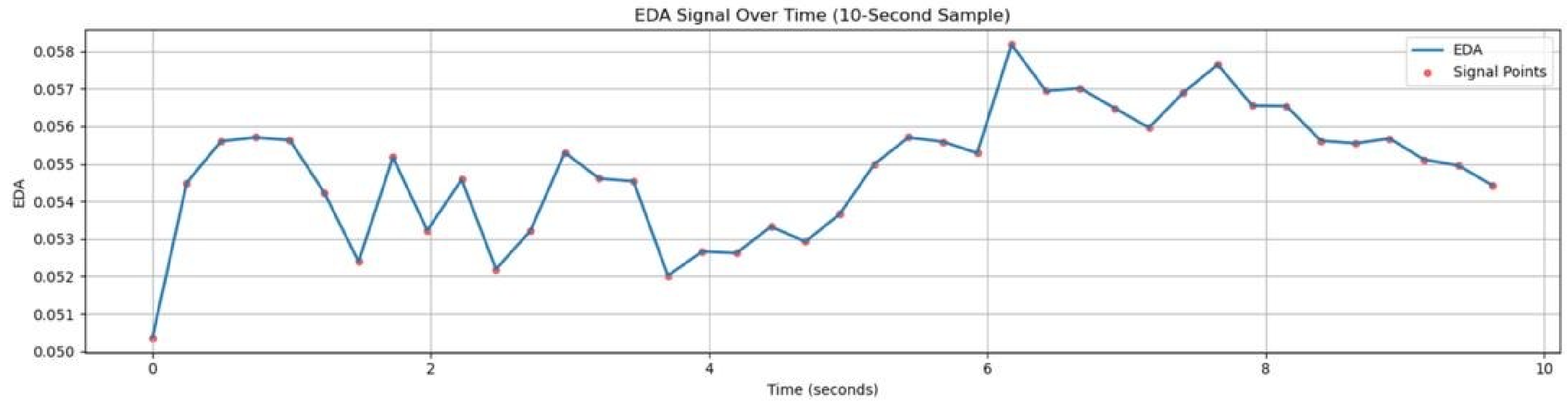
AnomalyDetectionCo × | AnomalyDetectionED ×

Edit in VS Code | Compute: Serverless Spark Compute - Available | Language: PySpark (Python)

Session not started | Configure session

```
plt.scatter(sample_df['elapsed_seconds'], sample_df[METRIC], color='red', s=20, alpha=0.6, label='Signal Points')
plt.title(f'{METRIC_TITLE} Signal Over Time (10-Second Sample)')
plt.xlabel('Time (seconds)')
plt.ylabel(METRIC_TITLE)
plt.legend()
plt.tight_layout()
plt.grid(True)
plt.show()
```

[9] ✓ - Command executed in 48 sec 81 ms by mlarcaro on 12:40:44 PM, 7/22/25



The figure above presents the Electrodermal Activity (EDA) signal over a **10-second window**, sampled at approximately 4 Hz, yielding 40 data points. The signal exhibits distinct temporal variation consistent with **low-level tonic and mild phasic electrodermal responses**. The temporal behavior can be decomposed into three annotated regions, which are analyzed below.

Region I (~0 to ~2.5 seconds): Baseline Shift and Initial Rise

In this initial segment, the EDA signal exhibits a **moderate increase in amplitude**, likely reflecting either the tail end of a preceding phasic response or a spontaneous increase in sympathetic arousal.

Let $x(t)$ represent the EDA signal at time t. Over this interval:

$$\frac{dx}{dt} > 0 \quad \text{and} \quad \mathrm{Var}(x_{[t_1,t_2]}) > \epsilon_1$$

with ϵ_1 a small but non-negligible variance threshold. The trajectory is smooth and monotonic, suggesting a physiological origin rather than a sensor artifact. However, the **absence of a sharp inflection** indicates this is more likely a **tonic shift** than a discrete phasic event.

Region II (~2.5 to ~6.5 seconds): Low-Amplitude Fluctuations

Proposed AI Technologies

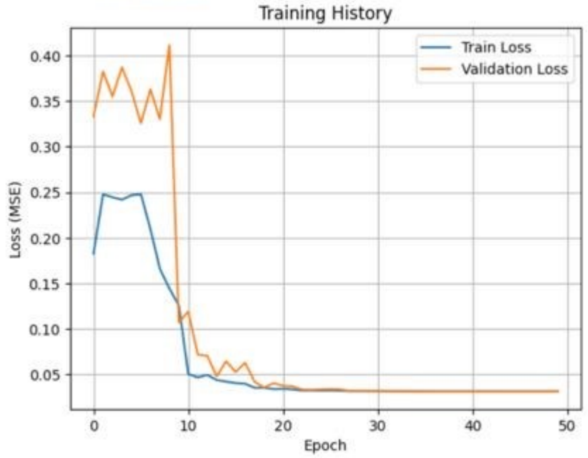
Microsoft Foundry | Azure Machine Learning

BioGen > biogen-nonprod-aml > Notebooks

AnomalyDetectionCo × | AnomalyDetectionED ×

Edit in VS Code | Compute: Serverless Spark Compute - Available | Language: PySpark (Python)

Session not started | Configure session



6.1.5 Final Training Summary

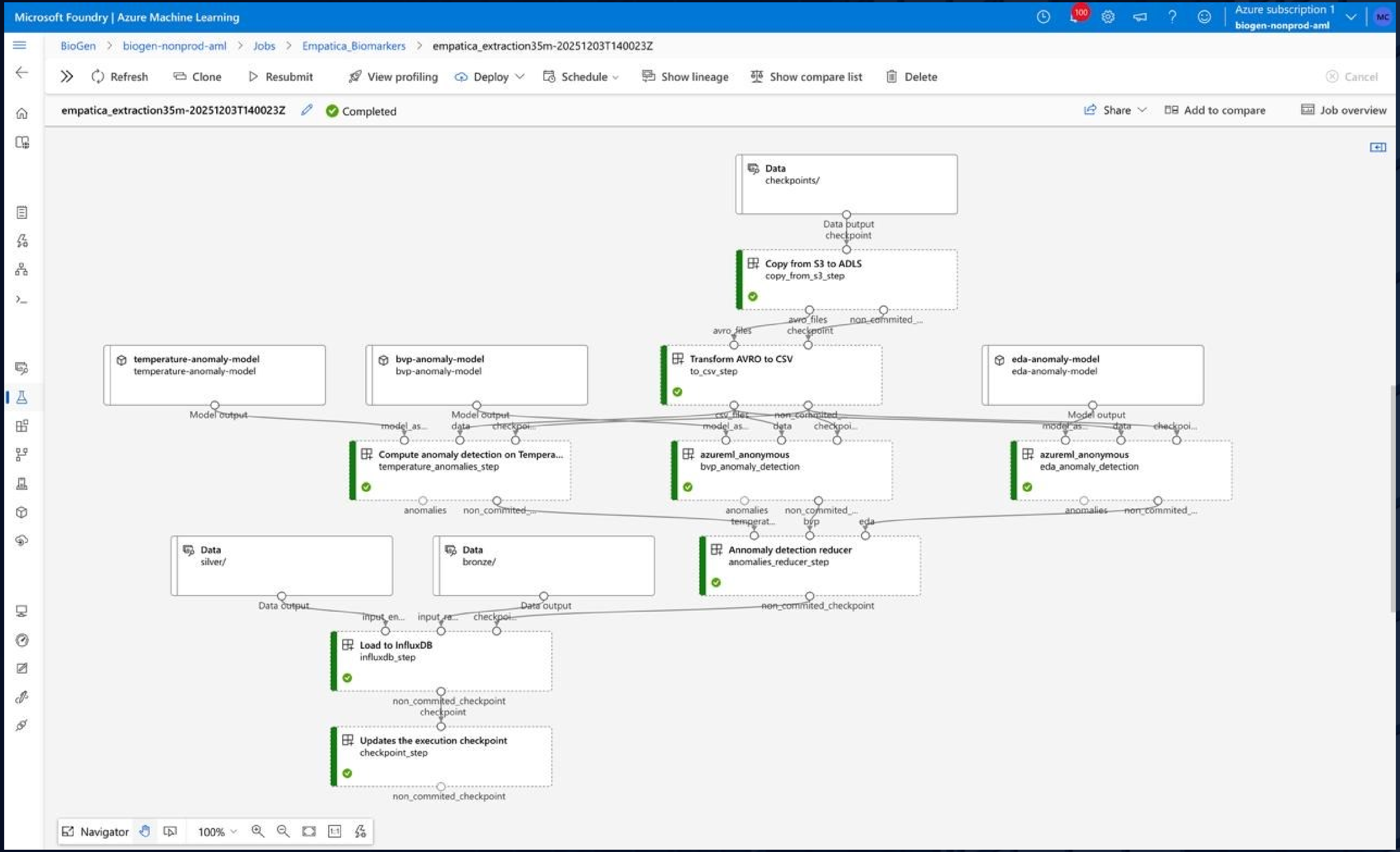
Total Epochs Trained: 50

- **Best Validation Loss (MSE):** ~0.032 (epoch ~28)
- **Training Loss at Best Epoch:** ~0.030
- **Learning Rate Schedule:**
  - Initial: 0.001
  - Reduced to 0.0005 after plateau in validation loss (~epoch 10)
  - Further reductions applied during later fine-tuning stages

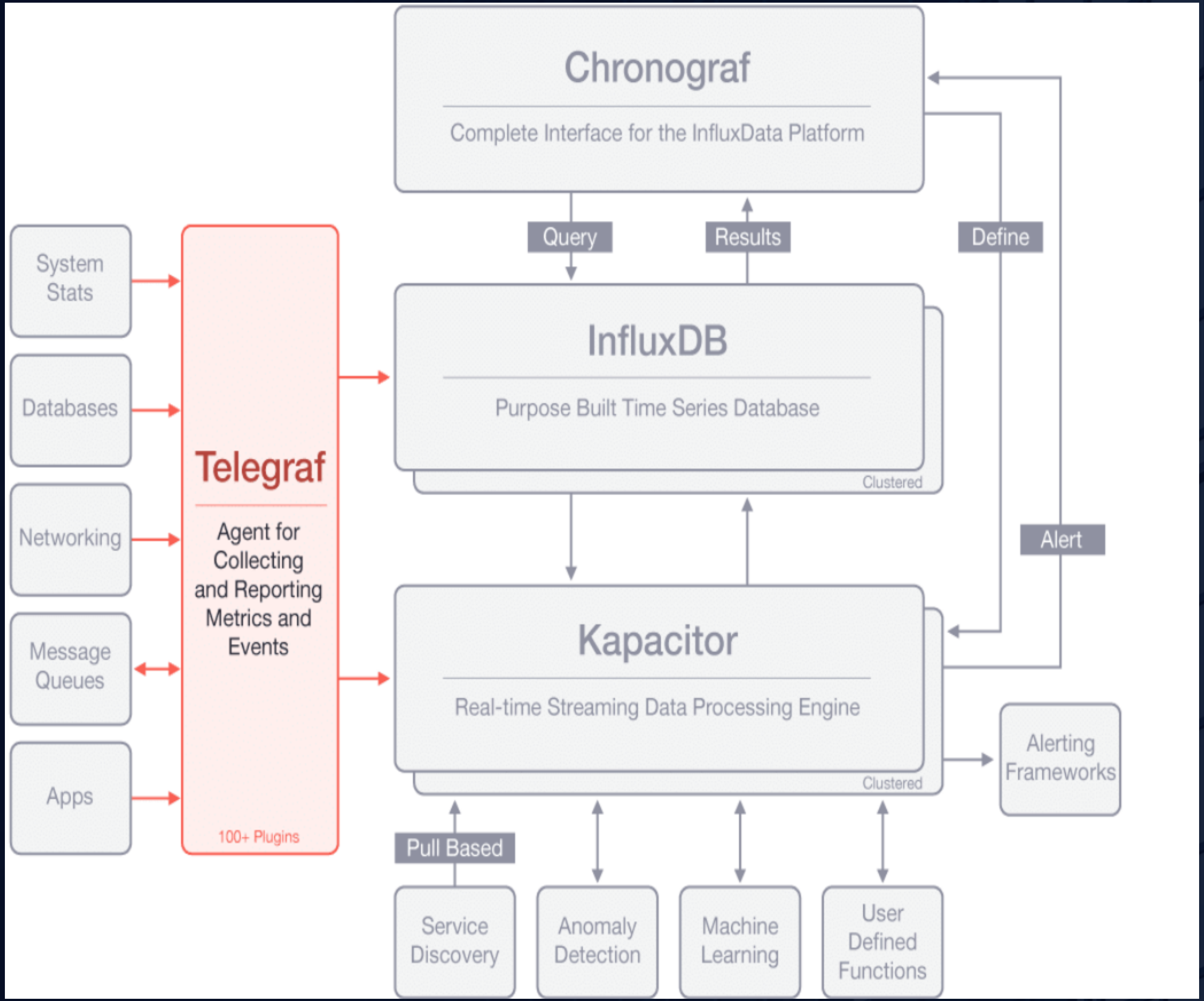
Metric	Observation	Meaning
Train loss	Smooth, consistent decrease to ~0.03	Model fits training data well
Validation loss	Drops sharply early, stabilizes around 0.032	Model converges quickly and maintains generalization
Learning rate decay	Triggered after plateau periods	Helped refine weights without large jumps

The results indicate the model achieves stable generalization within the first 25-30 epochs. Training beyond this range does not yield significant improvements, so future runs could use early stopping to save compute time.

Proposed AI Technologies



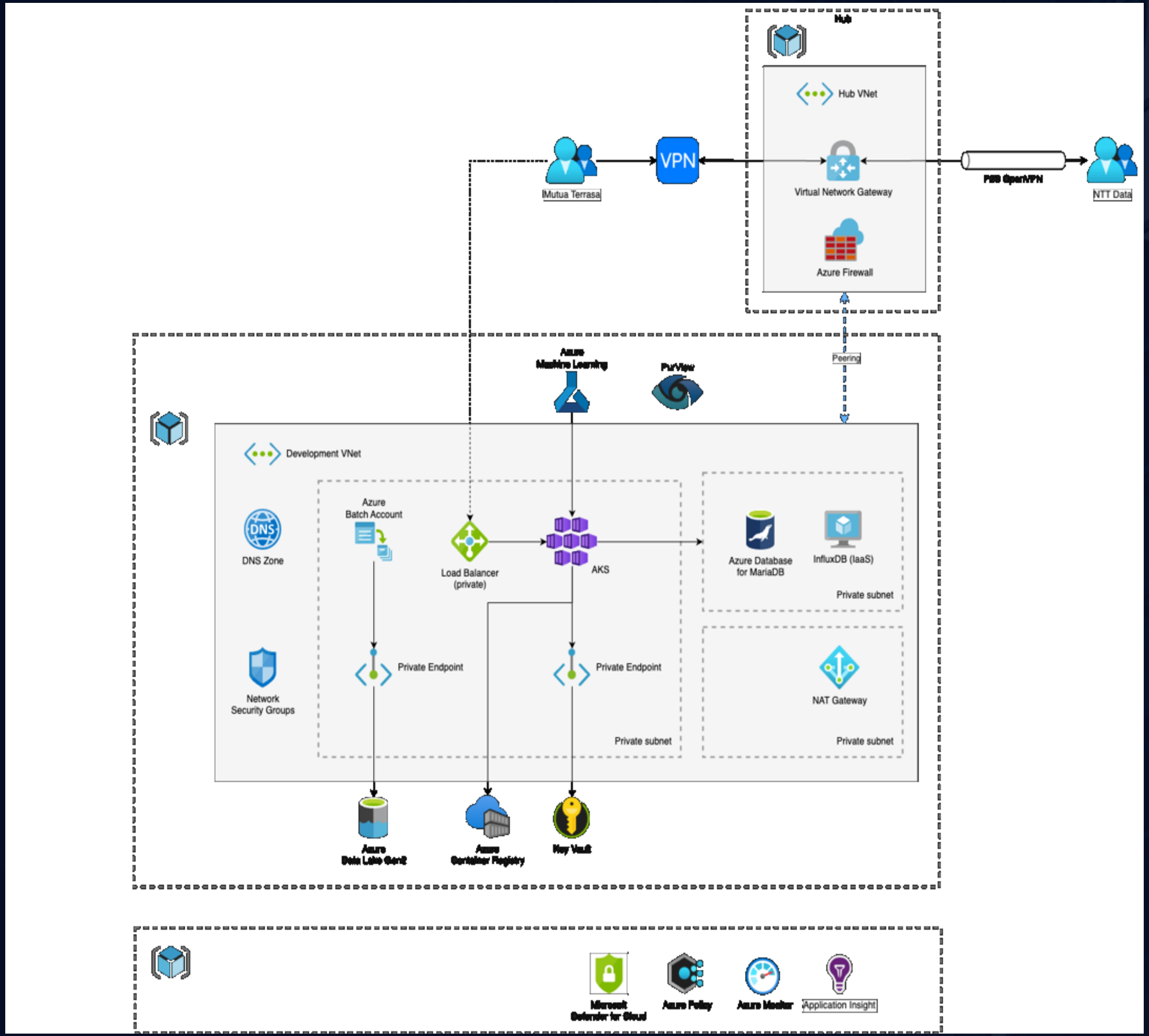
Proposed Big Data Technologies



InfluxDB is an open-source time-series database written in Go by InfluxData. It is optimized for high availability data retrieval, fast time-series data storage, and real-time analytics in fields such as operations monitoring, application metrics, IoT sensor data, and more.

Proposed Cloud Technologies

NTT DATA



Azure Architecture