BORACLE - Intelligent Algorithms (IA): AI-Driven Health & Performance Modeling from Wearable Devices

ARCS
AUTONOMY RESEARCH CENTER FOR STEAHM

Improving Social Medicine with Autonomy

Brandon Ismalej, Neha Ananthavaram, Abhinav Neelam, Zachary Frappier

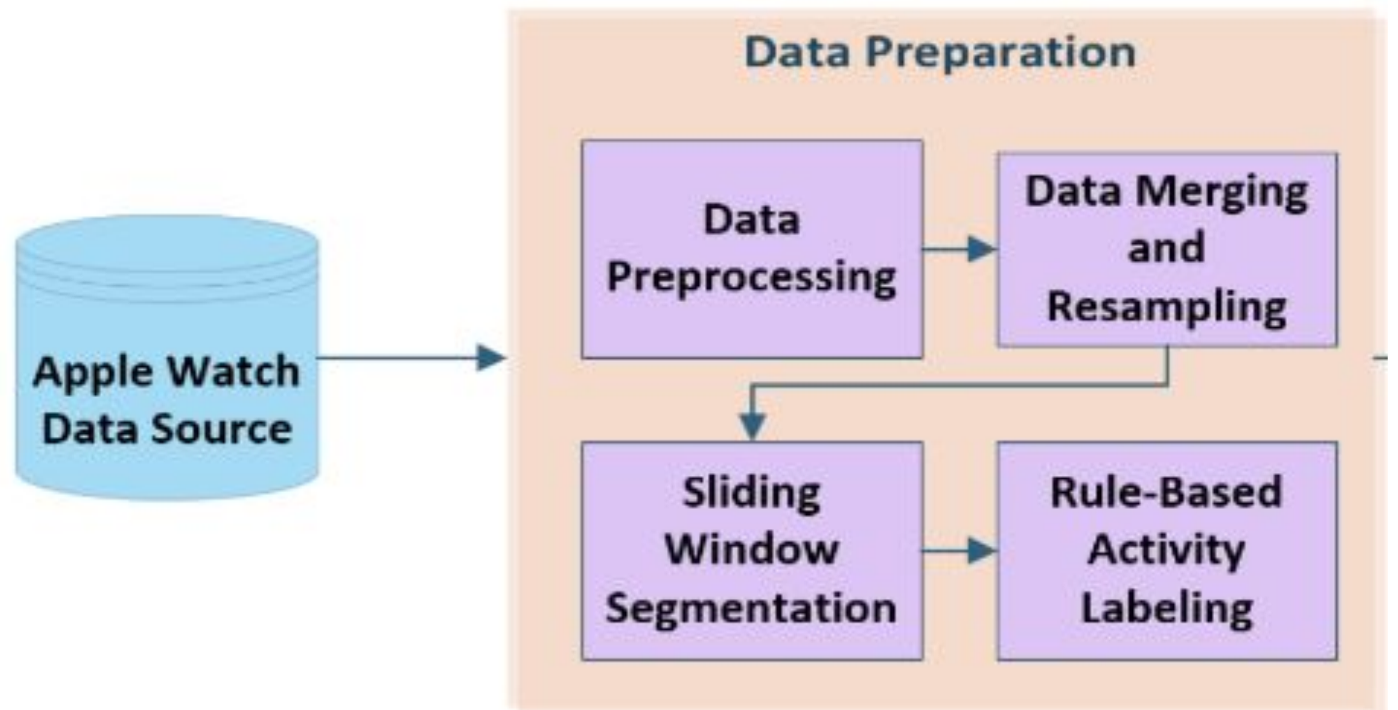

Figure 1: Pipeline for Activity Classification using Apple Watch Data.

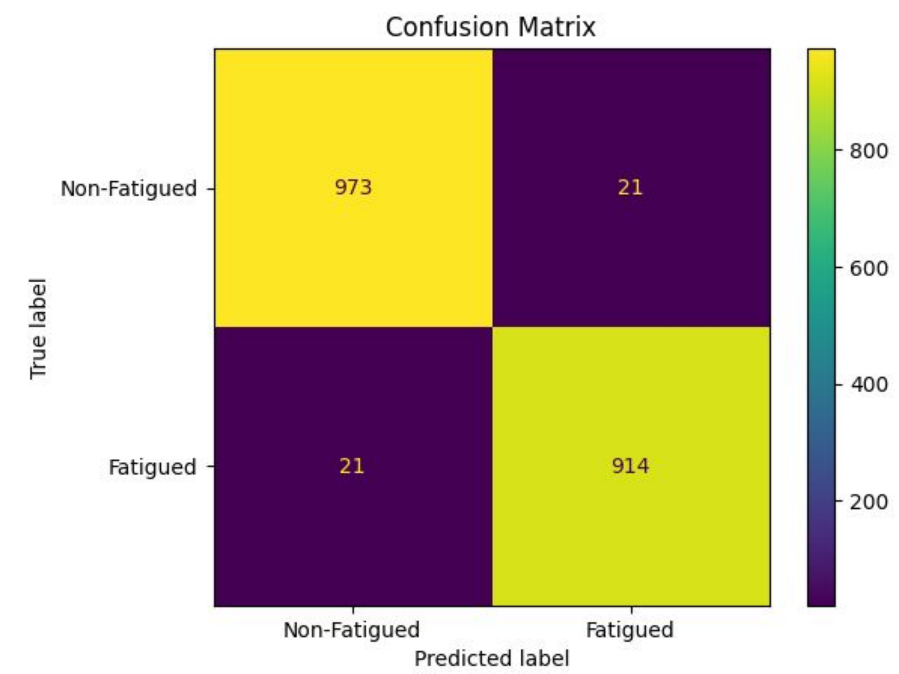

Figure 3: Confusion Matrix of Fatigue Prediction

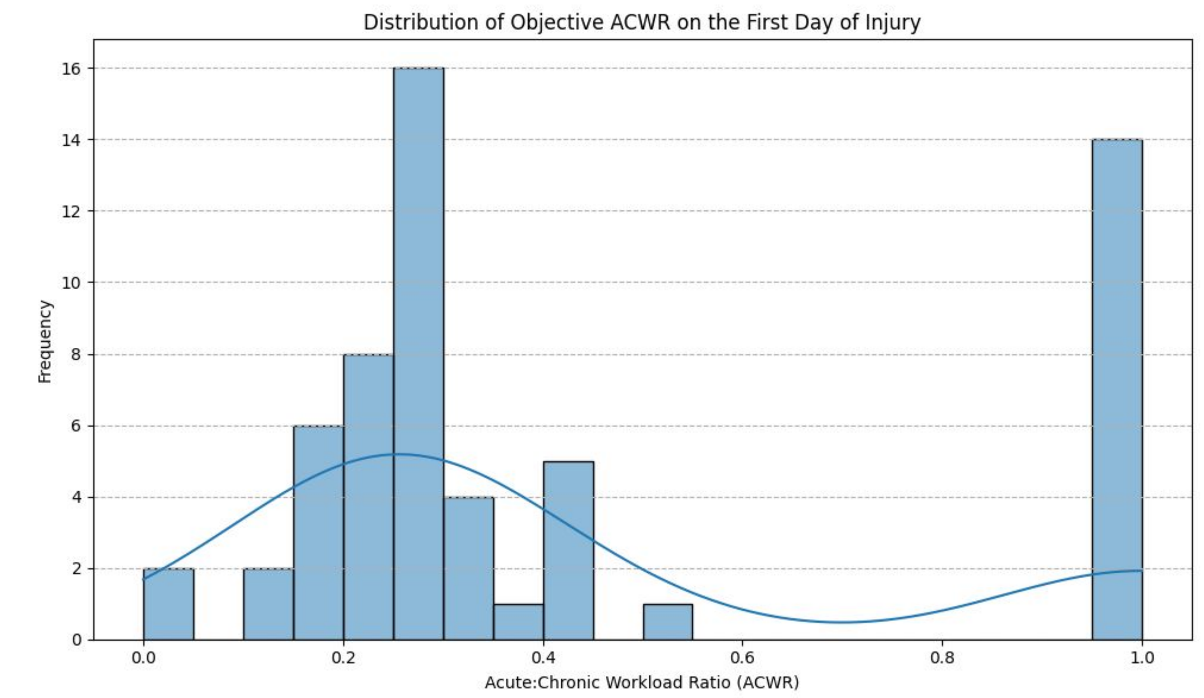

Figure 4: Distribution of ACWR of 1st Day of Runner Injury

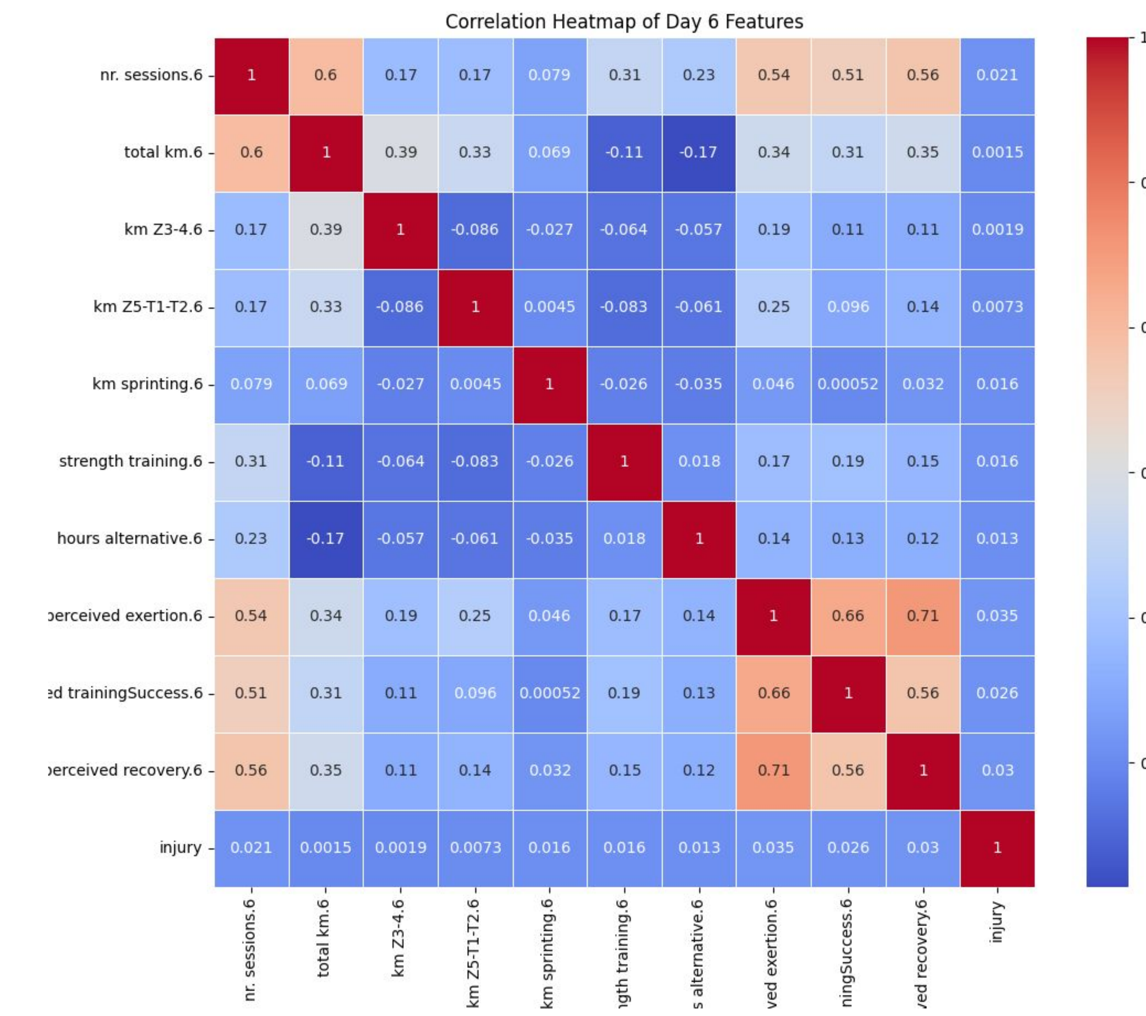

Figure 6: Correlation Matrix for Runner Injury Dataset Features

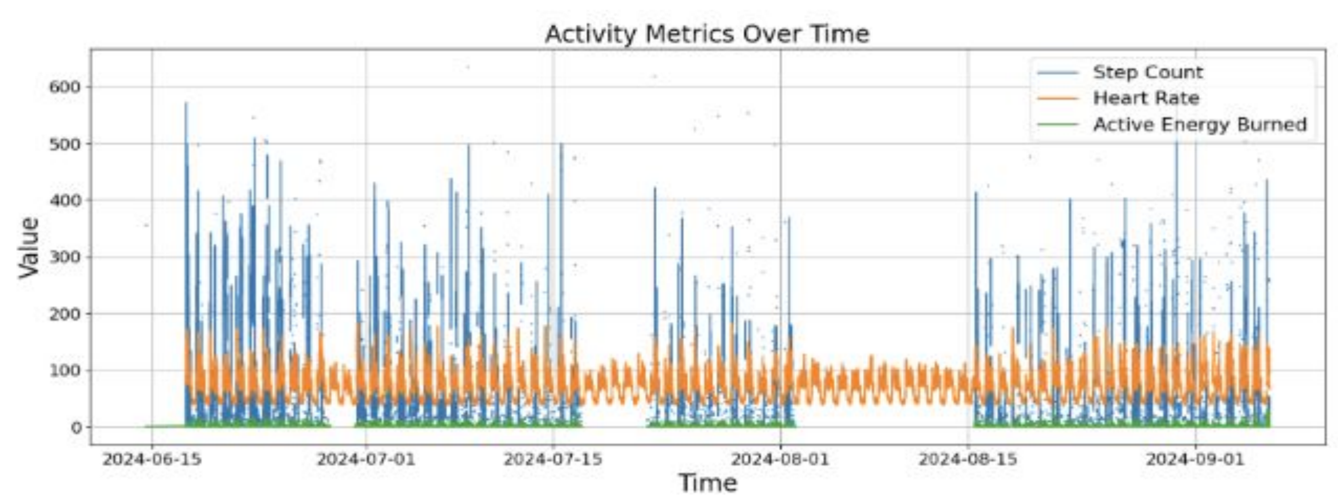

Figure 2: Time-series trends in step count, heart rate, and energy burned across multiple days

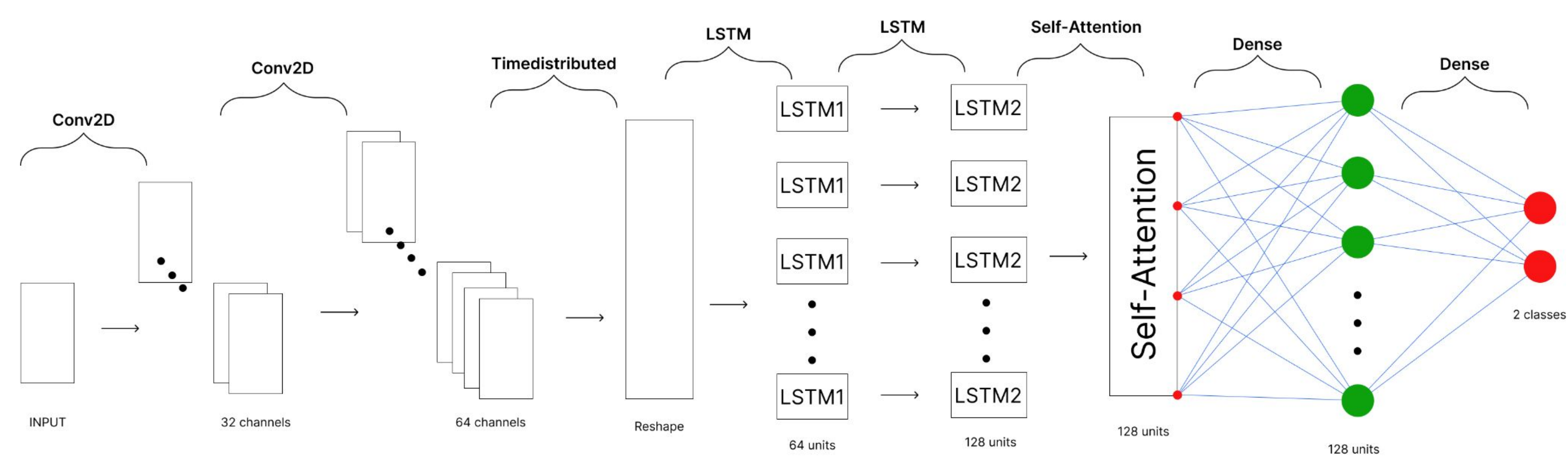

Figure 5: DeepConvLSTM with Self-Attention

Introduction

- **Objectives:** Develop Intelligent Algorithms to analyze and detect healthcare conditions using various machine learning (ML) models
- **Current Research Topics**:
  - Physical Activity Classification
  - Runner Injury Forecasting
  - Fatigue Prediction
- **Prior Research Topics**:
  - Detecting Heartbeat Irregularities (Arrhythmia) using ECG Data
  - Sleep Stage Classification and Stress Level Detection
  - Injury Forecasting for Soccer Athletes

Data Collection

- **Apple Watch (Series 7):** 24/7 data collection over 3 months (Jun. - Sept. 2024)
- **Metrics:** Steps, HR, Energy Burned, VO_2 Max, HRV
- Three IA team members

Public Datasets

- **Competitive Runners Injury Dataset [1]:**
  - 77 high-level middle and long distance runners, over seven years
  - Objective data from GPS watch, subjective data for exertion and success of training
- **DUO-GAIT Dataset [2]**
  - Public Gait Dataset for walking under Fatigue and Control Conditions
  - Provides IMU sensor data from 18 participants with 6 minute walks

Physical Activity Classification

- **Granularity:** Data was preprocessed, cleaned, and resampled into 5-minute sliding windows, resulting in 358 labeled segments; sample in Fig. 2
- **Labeling Approach – Rule-based logic** using heart rate and step count thresholds was used to label windows as:
  - Rest (e.g., HR < 70 bpm, 0 steps)
  - Walk (e.g., HR 70–110 bpm, 30–80 steps)
  - Run (e.g., HR 120–160 bpm, high step count & energy)
  - Cycle (e.g., HR 100–140 bpm, low step count, distance involved)
- **Results:**

Model	Accuracy	Macro F1-Score
Random Forest	96%	0.78
Support Vector Machine	92%	0.53

Runner Injury Forecasting

- **Exploratory Data Analysis insights** in Fig. 6:
  - Target feature imbalance: injury vs. no injury
  - Low feature correlation
- **Feature Engineering Metrics**
  - Training Load = Session Duration $\times$ Rate of Perceived Exertion
  - Training Strain = Σ training load $\times$ training monotony
  - Training Monotony = Avg. Training Load / Std. Deviation of Training Load
  - Acute Chronic Workload Ratio (Fig. 4): Acute Chronic Workload / Chronic Workload
- **Data Preprocessed**
  - Time-series; Sequences of 7 days
- **Neural Network Training:** Gated Recurrent Unit

Fatigue Prediction

- **DUO-GAIT** provides ACC + GYRO readings across various body locations
  - Utilized the left foot sensor located above the foot instep
- **Preprocessing:**
  - 3rd order Butterworth filter to smooth out sensor noise with 10 Hz cutoff freq
  - Padded stride windows to consistent length of 150 samples
  - Standard scale each feature channel independently based on training set
- **Custom DeepConvLSTM Architecture with Self-Attention Mechanism**
  - Single Head Dot-Product Self-Attention to identify key time steps, shown in Fig. 5
- **Supervised Learning Results**
  - 91% F1-Score using 1D-CNN Architecture
  - 97% F1-Score using DeepConvLSTM with Self-Attention Architecture
  - Class inference shown in Fig. 3

Future Exploration

- Expand exploration of diverse and representative datasets to enhance understanding of injury risk factors
- Continue training and fine-tuning deep learning architectures to achieve balanced performance - modest accuracy with strong generalization

References

[1] S. S. Lövdal, R. J. R. Den Hartigh, and G. Azzopardi, "Injury Prediction in Competitive Runners With Machine Learning," International Journal of Sports Physiology and Performance, vol. 16, no. 10, pp. 1522–1531, 2021, doi: https://doi.org/10.1123/ijspp.2020-0518.
[2] L. Zhou, E. Fisher, C. M. Brahms, U. Granacher, and B. Arnrich, "DUO-GAIT: A gait dataset for walking under dual-task and fatigue conditions with inertial measurement units," PubMed, https://pubmed.ncbi.nlm.nih.gov/37604913/