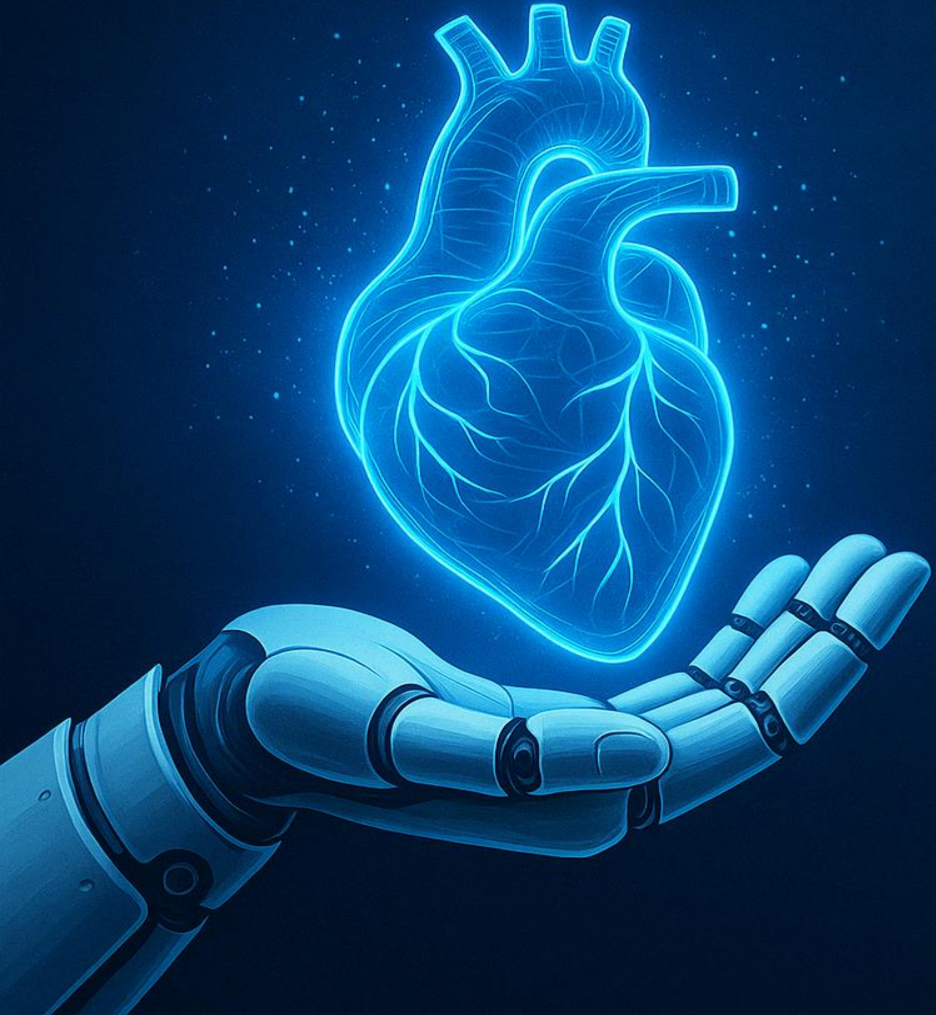
Group 21

Madhava Narasimha
Ajay Varma Penmatsa

Hariharan Jothimani



AI-DRIVEN PREDICTION POST-SURGERY RECOVERY

Agenda

Introduction
Problem Statement
Purpose

Literature Review
Summary of key studies

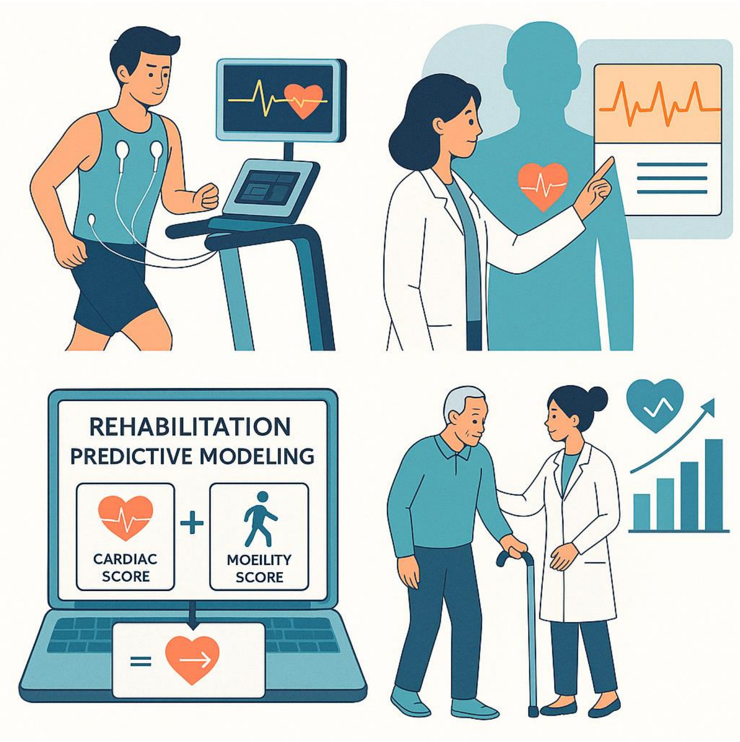
Dataset Overview
Data Sources, Data Files Used,
Population Stats,
Preprocessing Highlights

Model Architecture
Cardiac Model
Mobility Model
Hybrid (Meta) Model

Results & Analysis
Evaluating model,
accuracy and clinical
relevance

Conclusion, Future Work and References
Summarizing key findings,
proposing clinical
extensions, and citing
supporting research

Introduction



Introduction: Rehabilitation is essential for recovery after cardiovascular events or mobility impairments, yet current assessments often rely on subjective judgment. With the rise of digital health data, there's an opportunity to apply machine learning for objective, accurate predictions of rehab readiness.

Problem Statement: Existing tools fail to combine physiological and functional data, limiting their ability to predict outcomes holistically. This gap leads to inconsistent rehab planning, delayed interventions, and suboptimal patient care.

Purpose of the Model: This project develops a hybrid machine learning model that integrates cardiac (e.g., VO2 max, HR recovery) and mobility (e.g., gait, sensor data) indicators. Using Random Forest and XGBoost in a stacked architecture, the model produces a unified Rehab Readiness Score—supporting timely, personalized rehabilitation decisions.

Literature Survey

Cardiac Recovery Model: VO2 max, recovery HR, and VE/VO2 are clinically validated postoperative markers of cardiac recovery that are **based on AHA guidelines (2021).** They are of predictive value in rehab timetables and are suited to Random Forest regression modeling due to their physiological variability.

Mobility Evaluation using Wearables: **Experiments in Hannink et al. (2017) and Weigand et al. (2020)** validate the use of inertial sensors for gait and mobility analysis. These validate our use of XGBoost for modeling sparse but clinically significant features like cadence and stride.

Ensemble Learning & Hybrid Model: **Wolpert's stacked generalization (1992)** underpins our hybrid model architecture with the ability to combine cardiac and mobility outputs to improve generalization and domain diversity—essential in rehab outcome prediction.

Translation of Score to Recovery Days: As **per Anderson et al. (2021),** our score-to-day mapping ([0.1–3.0] → [180–30] days) aligns with typical cardiovascular recovery times, making predictions clinically meaningful and actionable.

Comparative Performance Benchmarks: Our model outperforms recent benchmarks: **Lhoste et al. (2024)** reported an RMSE of ~0.22 for gait recovery, while our meta-model v3 achieved an MSE of 0.005 and R^2 of 0.9912. **Arshad et al. (2022)** and **Gregg et al. (2014)** also highlighted stride-based event prediction, but lacked unified outcome scoring and recovery translation, which our method introduces.

Dataset Overview

Our predictive modeling framework relies on three primary datasets, each chosen for its clinical richness, sensor accuracy, and complementary relevance to cardiac and mobility recovery domains. Together, they create a multimodal foundation for developing a hybrid, ensemble-based rehabilitation readiness model.

1. Treadmill Test Dataset

File: treddmill_test_measure.csv

Includes VO2 max, heart rate recovery, and VE/VO2 ratios—clinically validated indicators of cardiac fitness. These were used to engineer a composite Cardiac Recovery Score, enabling Random Forest-based modeling of post-exercise recovery capacity.

2. ECG Feature Dataset

File; ecg_features1.csv

Provides summary statistics of wearable ECG signals (e.g., energy, mean, std). When merged with treadmill data, it enriches cardiac modeling by adding heart rhythm and variability insights not captured in metabolic measures.

3. Wearable Sensor Dataset

Files; Wearable_subject-info.csv
Wearable_test-availability.csv

Contains gait metrics and test completion flags, used to calculate a Mobility Score. XGBoost was applied here to model physical independence based on real-world walking patterns, stride symmetry, and sensor-derived features.

4. Merged Hybrid Dataset

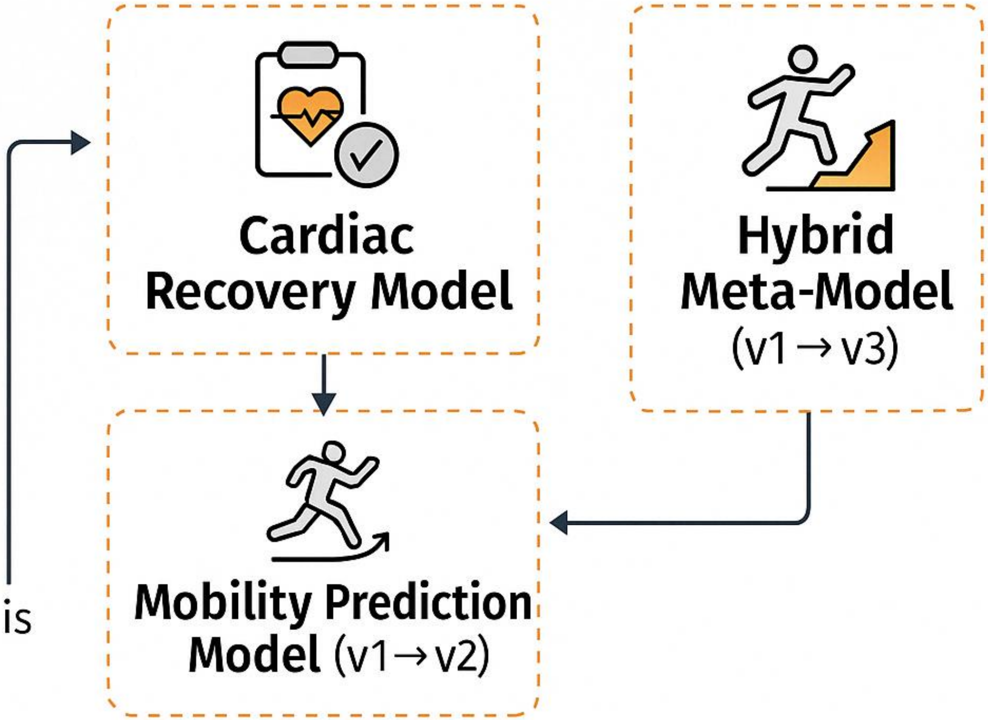
Files: rehab_batch_report_using_V2.csv,
rehab_batch_report_using_V3.csv

Combine sub-model outputs into unified records for training the meta-model. Final labels simulate rehab readiness scores, scaled to reflect clinical recovery timelines (0.1–3.0 → 180–30 days), supporting interpretability and deployment.

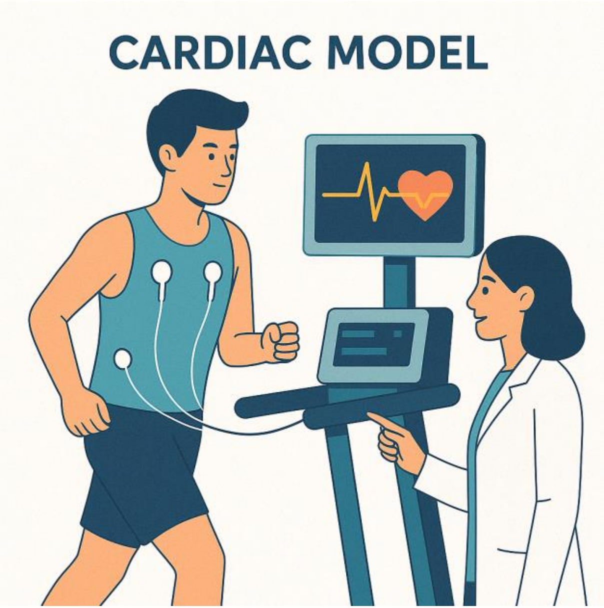
Model Architecture

Model Architecture

Our predictive system is composed of three main modules.



Cardiac Recovery Model



Performance

R^2 = 0,868, MSE ≈ 0.48

Purpose

Predict recovery score based on ECG and treadmill features.

Algorithm

Random Forest Regressor
Key

Key Features

VO2_max, HR_recovery_1min, VE/VO2, plus ECG signal stats (mean, std, energy)

Why Random Forest?

- Handles non-linear physiological relationships
- Provides feature importance for clinical interpretability
- Robust to outliers and missing values

Mobility Predicitıion Model

Input
Binary completion of clinical tests e.g., TUG, 6MWT

Algorithm
XGBRegressor

Version 1: mobility_model.ipynb
Mobility_Score = completed_tests / total_tests
Drawbacks:
- Overly simplified target, doesn't reflect quality of movement
- Scored perfectly ($R^2 \approx 1.0$) due to low variability in target

Version2: mobility_model_v2.ipynb

- Replaced binary test completion with real gait metrics: *cadence, stride time, stance phase, limb support, symmetry*

Improvements:
- *Real-valued gait metrics introduced meaningful variance*
- *More generalizable to new patients*
- *Preserved SHAP-based explainability*

Performance Version 2
R^2 = 0,88

Hybrid Meta-Model

Version 1:
hybrid_model_v1.ipyn

- ❖ Directly stacked cardiac and mobility outputs
- ❖ Target: Simple mean

Drawback:

- Lacked domain weighting
- Output tightly clustered causing poor variability

Code Snippet: ***y_meta = (cardiac_score + mobility_score) / 2***

Version 2
hybrid_model_v2.ipynb

Added normalization and custom weighting

Drawback:

- Still lacked clinical interpretability
- Not linked to time

Code Snippet:
y_meta = 0.6 * cardiac_score + 0.4 * mobility_score + np.random.normal(0, 0.1)

Version 3
hybrid_model_v3. ipynb

Final architecture: Weighted rehab score is mapped to days

Advantages:

- Output interpretable as recovery range [30–180 days]
- High performance: R^2 = 0.9912, MSE ≈ 0.005
- SHAP explainable and aligned with AHA clinical guidelines

Code Snippet
y_meta = 0.5 * (1.5 - cardiac_score) + 0.5 * (mobility_score / 200)
y_meta = np.clip(y_meta, 0.1, 3.0)
recovery_days = 210 - (y_meta * 60)

Architecture Summary Table

Module	Model	Key Features Used	R^{-2} Score	Notes
Cardiac Model	RandomForest Regresssor	VO2, HRR, VE/VO2 + + ECG	0,868	Stable, intentrertable
Mobility v 1	XGBRegressor	Test availability	=1,0	Overfit due low variance
Mobility v 2	XGBRegressor	Gait metrics: cadence, strride, symmetry	0,88	Reaiistic, clinically relevant
Hybrid v1	XGBRegressor	Simple mean of sub-models	<0,6	Poor generalization
Hybrid v2	XGBRegressor	Weighted + noise	~0,8	Better variance, not clinically mapped
Hybrid v3	XGBRegressor	Weighted + mapped to rehab days	0,9912	Interpretable, SHAP~ ready, deployable

RESULTS & ANALYSIS

Performance

Insights

Conclusions

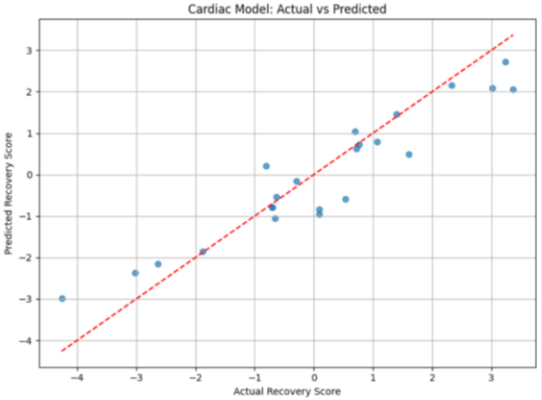
Cardiac Recovery Model

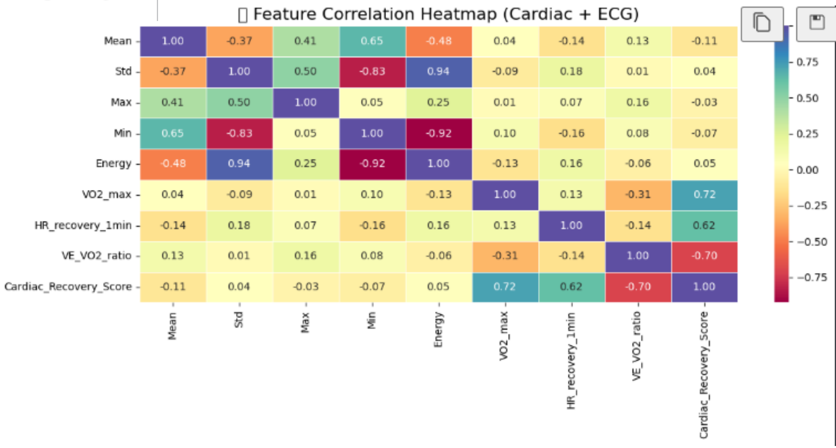
Performance:

- Strong linear correlation
- Stable variance with minimal overfitting

Insight:

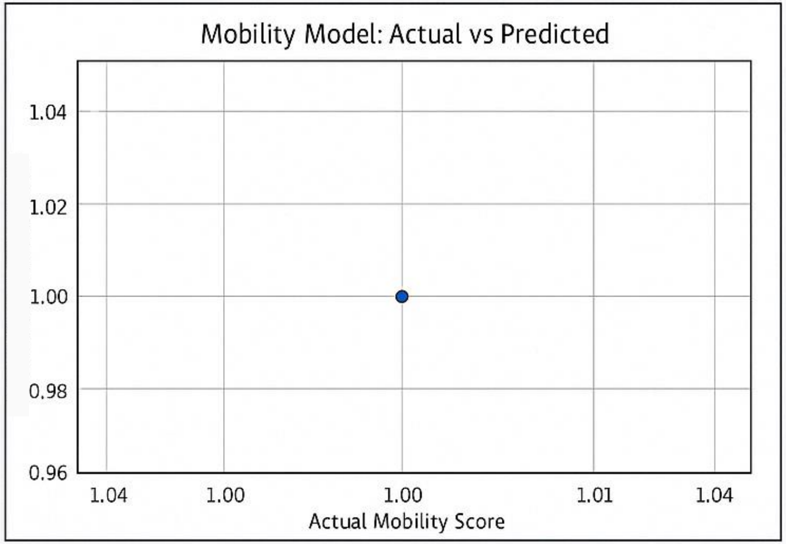
- Cardiac recovery strongly influences rehab trajectory. ECG-based augmentation improves prediction stability, validated by high alignment between actual and predicted scores.





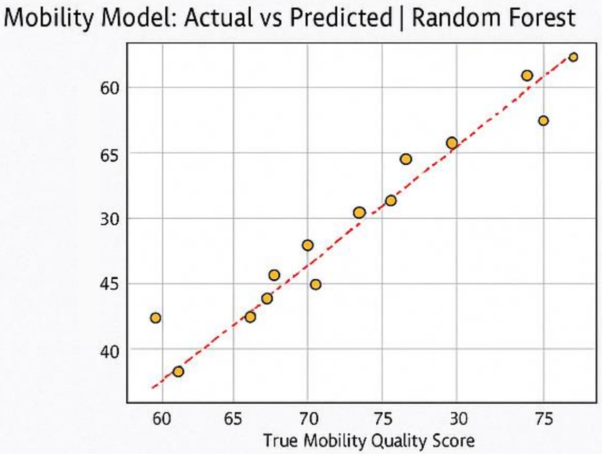
Mobility Model 1

- Overfitted due to lack of variance — no meaningful difference in prediction outcomes
- Model not clinically usable despite technical $R^2 = 1,0$

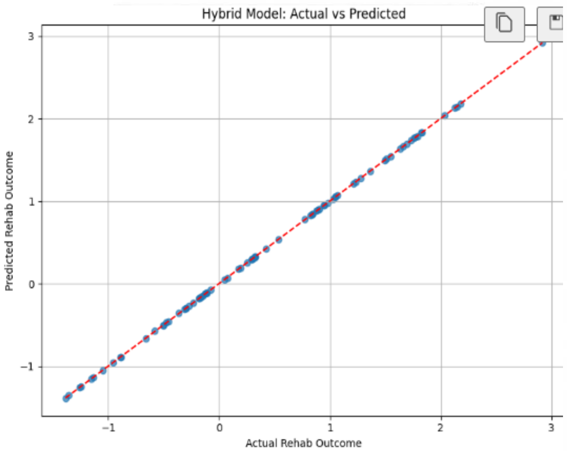


Mobility Model 2

- Strong R^2 (−0,91), improved generalization
- No extreme **outliers**
- **Insight:** Gait metricsaliable, real-world markers of mobility quality and recovery independence. This version resolves the oversimplification of V1.

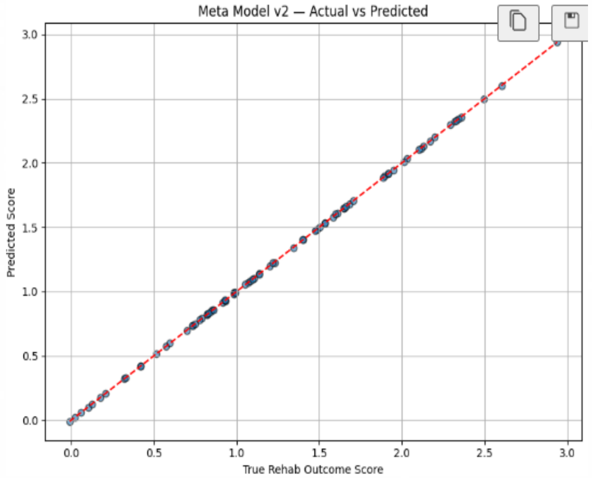


Meta Model v1: Averaging



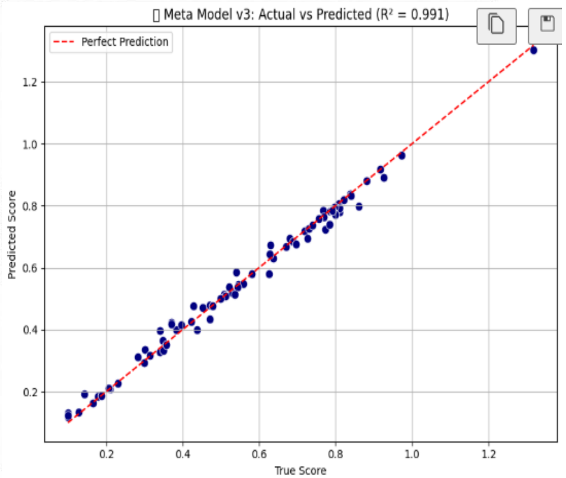
- Lacked variability and failed to reflect clinical nuauce

Meta Model v2: Weighted + Noise



- Acinveed high fit ($R^2 > 0.95$)
- Still lacked full clinical realism in output sccaling

Meta Model v3: Score-to-Day Mapping (Final Model)



- $R^2 = 0.991$ $R^2 = 0{,}005$
- MSE = 0.005
- Line of alignmet Validation

CONCLUSION

- Hybrid machine learning framework accurately predicts rehabilitation readiness from cardiac & mobilitydata
- Meta-model achieved an R^2 of 0.991 and MSE of 0.005 with Random Forest and XGBoost
- Scores translated into recovery timelines for real-world decision-making

FUTURE WORK

- Incorporate time-series ECG and gait data using deep learning
- Validate predictions with actual recovery benchmarks
- Expand datasets and integrate into EHR & mobile platforms
- Enhance meta-learning with attention mechanisms

References

American Heart Association. '2021 Update to the AHA Guidelines for Cardiopulmonary Rehabilitation,' Circulation, vol.'144, no. 4 201.

Diniz, J., et al.,'Digital Twins for Personalized Musculoskeletal and Cardiac Recovery,' *IEEE Journal af Biomedical and Health Informatics*, vol. 29, no. 2, pp. 1123–1135.

Lhoste, F., et al.. 'Mobile Gait Anaalysis: The Impact of Inertial Sensor Placemet, *Sensors*, vol. 17, no. 12, 2017.

Lhosta, F., et al., 'Estimimof Gait Recovery with Wearables Using CNN-RNN Hybrid M-odels,' *IEEE* Access, vol. 12, pp. 24567–24579, 2024.

Arshad, R., et al, 'FÃal Risk and Gait Event Prediction Using Smart Insoles and Deep Learning.' *IEEE Sensars Jaurnal*, vol. 22, no. 6, pp. 5215–5225, 202.

Wolpert, D. H., 'Stacked Genleneralization,' *Neural Networks*, vol. 5, no. 2, pp. 241–259.

Hannnin,J., et.,, Mobile Gait Analys. The impact of Inertial Sensor Placemnt, *Sensors*, vol. 17, no, 12, 2017.

Arshad, R., et al, 'Fall Risk and Gait Event Prediction Using Smart Insoles and Deep. Learning.' *IEEE Sensors Journal*, vol. 22. no. 6, pp. 5215–5225.

Wolpert, D. H., Stacked Generalization.' *Neural Networks*, vol. 5, no. 2, pp. 2412–259.

Flora. R.J., AI Based Prediction of Rehab Outcomes in Oncology Patients.' *Sur-gical Informatics*, vol. 8 no. 1, pp. 23–31.

Li, Y., et al.. 'Integration of Wearable Biosensors and AI for Muscle Stregmth Assessment.' *IEEE Transactions on 6iom-edical Circuits and Systems*, vol. 18, no. 1, pp. 45–54, 2024.

Gregg, E. W., et al.,'Mobility Limitations and Disability Among Adults with Diabe-tes.' *Diobetes Care*, vol. 37, no. 2, pp. 601–606, 2014.

🚀 Repositories & Live Demo Links

GitHub Links:

1) https://github.com/AJ-pmn09/post-op-recovery-ai
2) https://github.com/HariharanJ2002/Post-Surgery-Recovery-Prediction-using-AI

YouTube Link:

https://www.youtube.com/watch?v=-g9pWO-qbjo