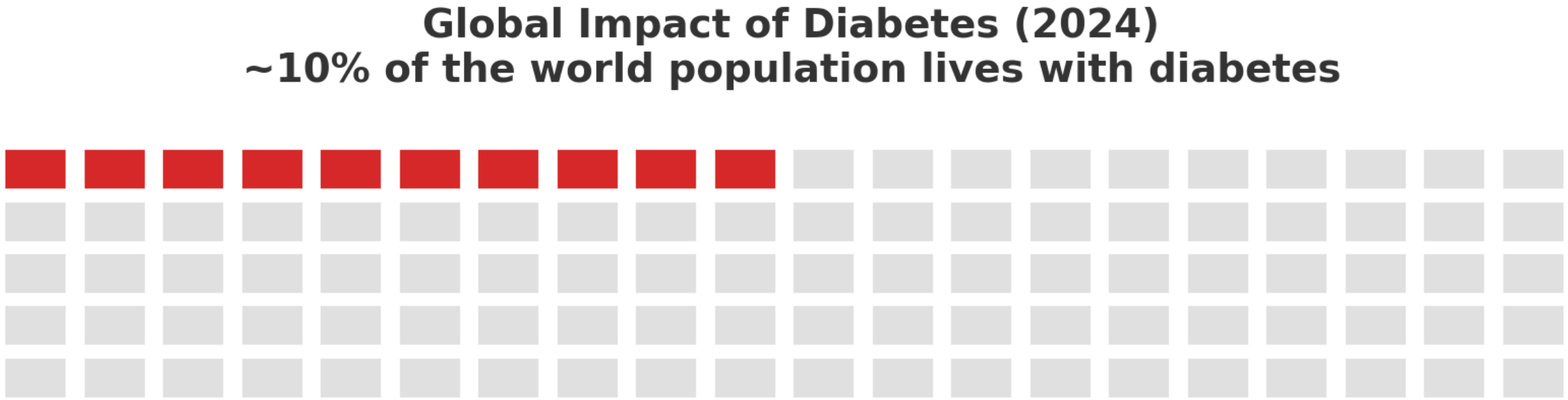
Diabetes Management:
Non-Invasive Glucose Monitoring via ECG and Machine Learning

ITADATA 2025:
The 4th Italian Conference on Big Data and Data Science,
9-11 September 2025,
Turin, Italy.

DIABETES

Diabetes affects [1] over 830 million people worldwide.



[1] https://www.cdc.gov/diabetes/php/data-research/index.html

DIABETES

Diabetes affects [1] over 830 million people worldwide.

There are **two** primary types of diabetes (type 1 and type 2).

Diabetes can lead to many different **complications** (including **death**) and can impact multiple organs.

We can handle it! **But...**

[1] https://www.cdc.gov/diabetes/php/data-research/index.html

DIABETES HANDLING

Managing Diabetes Requires Control of glucose and Insulin Levels.

Flash **G**lucose **M**onitoring (on demand!).

Continuous **G**lucose **M**onitoring (stream of data).

Drawbacks of FGM and CGM

Skin irritations.

Invasivity [2].

Costs.

[2] Skin Problems due to Treatment with Diabetes Technology: A Narrative Review (https://esmed.org/MRA/mra/article/view/4747)

Aim of our work

To propose a non-invasive tool to assist in blood glucose monitoring that can:

- **Be a «proxy»** (can estimate) for glucose level;
- Feel less invasive;
- Be cheaper than FGM and CGM;

Heath beat → Glucose Level

Some studies [3] [4] have suggested a possible correlation between ECG and glucose, hence our choice to use ECG as a possible indicator.

This study applies statistical methods and machine learning to examine the link between heart rate changes and blood glucose levels in diabetic patients, with the aim of exploring new non-invasive strategies for glucose monitoring.

[3] Identification of hypoglycemia and hyperglycemia in type 1 diabetic patients using ECG parameters (https://pubmed.ncbi.nlm.nih.gov/23366486/)
[4] Poor glycemic control is associated with increased diastolic blood pressure and heart rate in children with Type 1 diabetes (https://pubmed.ncbi.nlm.nih.gov/15207840/)

Research Questions

Research Questions 01

Is it possible to detect and differentiate between hypoglycemia and hyperglycemia events using ECG time series?

Research Questions 02

Is it possible to detect changes in a heartbeat and estimate HR trends using glucose monitoring data time-series?

HUPA DATASET [5]

Feature	Description
Glucose	Patient's blood glucose levels
Calories	Patient's caloric intake
Heart Rate	Patient's heart rate in that moment.
Glucose Categories	very low (<60), low (60–89), good (90–179), high (180–249), very high (>250)
Glucose Rate Moving Average	11-observation moving average (BGL)
Heart Rate Moving Average	11-observation moving average (HR)
Glucose Rate Left Shifted Average	5-observation left-shifted average (BGL)
Heart Rate Left Shifted Average	5-observation left-shifted average (HR)
Max in Window	Maximum BGL in sliding window
Frequency Above Threshold 1.15	Glucose values >1.15% of max in window
Frequency Above Threshold 1.25	Glucose values >1.25% of max in window
Min in Window	Minimum BGL in sliding window
Frequency Below Threshold 1.15	Glucose values <1.15% of min in window
Frequency Below Threshold 1.25	Glucose values <1.25% of min in window
Heart Mean (HM)	Average HR over six observations
Heart Trend (HT)	HR deviation from previous HM → U (increasing), D (stable/decreasing)

Original features

Engineered features

[5] https://data.mendeley.com/datasets/3hbcscwz44/1

RESULTS DURING TESTING

ID	Best Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
1	EBT	67.8%	66.69%	62.31%	64.43%	74.46%
2	EBT	66.7%	73.22%	68.36%	70.71%	72.79%
3	EBT	70.9%	77.69%	72.71%	75.12%	77.28%
4	EBT	79.0%	81.37%	91.67%	82.05%	81.37%
5	EBT	81.3%	91.18%	84.65%	87.79%	84.93%
6	EBT	73.8%	77.87%	73.54%	75.65%	81.25%
7	EBT	68.0%	73.57%	68.47%	70.93%	74.77%
8	EBT	69.2%	74.77%	70.93%	72.80%	76.94%
9	WKNN	73.8%	47.43%	60.52%	53.18%	75.10%
10	EBT	68.3%	72.52%	68.66%	70.54%	74.90%
11	WKNN	65.5%	69.65%	67.21%	68.41%	72.17%
12	WKNN	64.5%	65.67%	65.22%	65.45%	70.87%
13	EBT	68.4%	73.89%	68.88%	71.30%	74.99%
14	EBT	70.0%	75.32%	70.84%	73.01%	76.51%
15	EBT	66.8%	73.21%	68.61%	70.84%	72.79%
16	WKNN	67.1%	63.82%	64.87%	64.34%	73.81%
17	EBT	68.3%	74.45%	69.90%	72.10%	74.84%
18	EBT	67.5%	72.40%	67.26%	69.74%	73.94%
19	EBT	66.8%	71.22%	67.38%	69.25%	73.28%
20	EBT	66.7%	71.82%	67.41%	69.54%	73.43%
21	EBT	72.1%	76.58%	71.82%	74.13%	78.62%
22	EBT	71.8%	78.83%	74.74%	76.73%	78.21%
23	EBT	64.3%	70.05%	65.12%	67.50%	70.03%
24	EBT	60.4%	64.87%	60.25%	62.47%	64.91%

EBT = Ensemble Bagged Trees
WKNN = Weighted K-Nearest Neighbors

Merged Tukey and ANOVA Results for Each Patient

PID	F value	Pr(>F)	SC (%)	MeD	MaD	MiD
1	51.01	<2e-16	80%	-1.23	11.12 (VLVH)	-7.59 (VHH)
2	14.96	4.04e-12	70%	-0.98	8.28 (VLVH)	-6.56 (VHH)
3	84.81	<2e-16	90%	0.89	13.21 (VLG)	-10.38 (VHG)
4	49.62	<2e-16	75%	-0.45	7.00 (VLVH)	-5.88 (VHL)
5	14.41	1.11e-11	85%	1.11	12.74 (VLVH)	-7.06 (VHG)
6	12.66	3.43e-10	80%	-1.67	6.19 (VLVH)	-7.66 (VHL)
7	30.49	<2e-16	75%	0.55	8.40 (VHL)	-6.69 (VLH)
8	10.76	4.85e-07	50%	-0.34	3.18 (VHG)	-1.41 (GH)
9	15.57	1.29e-12	70%	0.23	8.27 (VLVH)	-6.12 (VLH)
10	52.34	<2e-16	90%	2.02	13.19 (VLVH)	-7.59 (LH)
11	55.83	<2e-16	80%	1.12	7.63 (VHL)	-7.80 (VHG)
12	12.19	7.55e-10	65%	0.78	6.45 (VLG)	-4.48 (GL)
13	73.36	<2e-16	75%	1.56	10.28 (VLG)	-9.65 (VLL)
14	20.89	<2e-16	70%	0.89	5.93 (VHL)	-5.74 (VHG)
15	73.02	<2e-16	85%	2.34	7.84 (VHL)	-7.84 (VHG)
16	45.41	<2e-16	60%	-1.34	11.07 (VLG)	-8.69 (VLL)
17	24.26	<2e-16	90%	0.87	11.12 (VHG)	-7.62 (VLL)
18	21.88	<2e-16	65%	-0.45	5.73 (VLG)	-6.55 (VHH)
19	24.29	<2e-16	80%	-1.12	6.65 (VHL)	-4.79 (VLL)
20	21.19	<2e-16	70%	0.43	5.99 (VHG)	-3.32 (LH)
21	26.22	<2e-16	75%	-0.56	7.98 (VHH)	-6.76 (LG)
22	2.04	0.106	–	–	–	–
23	403.2	<2e-16	85%	1.22	14.60 (VHL)	-13.89 (VHVL)
24	48.08	<2e-16	85%	1.22	14.60 (VHL)	-13.89 (VHVL)

The first two columns show the results of the **ANOVA test**, while the remaining columns show the results of the **Tukey test**.

Results: RQ1

- **Method:** ANOVA + Tukey post-hoc on 24 T1DM patients
- **Finding**: 23/24 patients → significant HR differences ($p<0.05$)
- **Conclusion:** HR variability robustly tracks BGL fluctuations
- **Key result:** Strongest effects at extreme BGL states

Results: RQ2

- **Feature:** Heart Trend (HT) → Rising (U) vs Stable/Decreasing (D)
- **Finding**: Moderate–high predictive accuracy across patients
- Best model: Ensemble Bagged Trees > others

Limitations

- **Small cohort (**24 T1DM patients, no healthy or T2DM controls**) → limited generalizability**
- 5-min sampling & consumer wearables **→ temporal gaps, noise**
- Empirical tuning of model parameters → **no sensitivity analysis**
- **Confounders** (sleep, BMI, stress, medication) **not integrated**

Future Work

- Larger & more diverse cohorts (incl. healthy controls, T2DM)
- Higher sampling frequency / multimodal data (ECG, PPG, accelerometry)
- Systematic parameter optimization + sensitivity analysis

Future Work

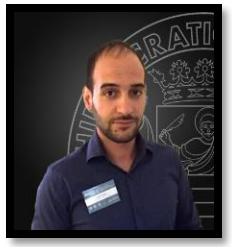
- Personalized & privacy-preserving modeling (Digital Twins, Federated Learning)

- Incorporate additional physiological/behavioral variables & apply explainability (LIME, SHAP)

Thank you

Matteo Ruggiero
Department of Computer Science
University Of Salerno

SCAN ME



Luigi Di Biasi, Fabiola De Marco, Alessia Auriemma Citarella, Genoveffa Tortora

Department of Computer Science
University of Salerno

ANY QUESTION?