Towards Smarter Home Spirometry: Leveraging Built-In Smartphone Sensors

Darcey Gardiner, Jacob Newman, Andrew Wilson, Min Hane Aung

University of East Anglia

Ethical approval for the collection and use of this dataset was granted by the University of East Anglia, Faculty of Science Research Ethics Subcommittee (reference ETH2425-1825).

1 INTRODUCTION

Spirometry measures lung function by assessing forced exhalation after maximal inhalation. During the COVID-19 pandemic, at-home spirometry became more prevalent, prompting studies comparing it to clinic-based testing. Results show at-home tests can be less accurate, especially without supervision, due to user error, poor training, or device limitations.

Our study explores the use of mobile phone sensors to detect and classify these errors by capturing additional data during at-home spirometry tests – potentially identifying errors that may not be evident in spirometry curves alone.

2 METHODS

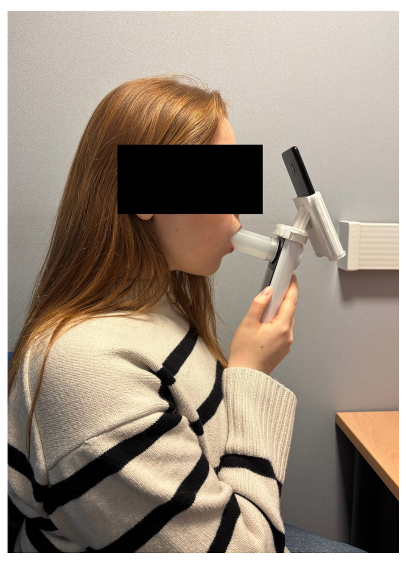

Figure 2. Participant using the attached device during testing

Ten healthy participants (5M/5F) used a spirometer paired with a mobile phone running a custom app to record microphone, accelerometer, and gyroscope data. Participants performed both normal efforts and simulated use-errors in a soundproof room.

A foamboard-mounted phone attachment was used to collect consistent sensor data (Fig. 1). Seen in Figure 2 during the recording.

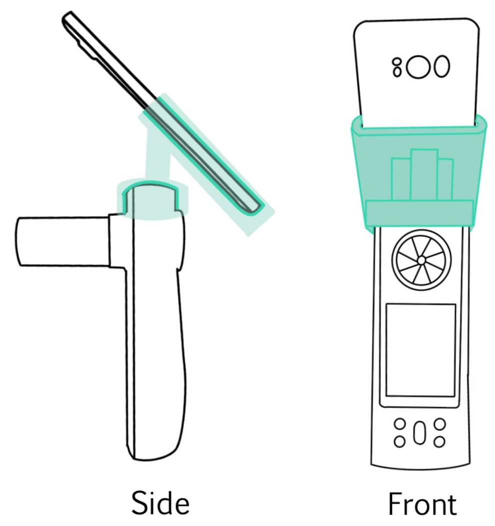

Figure 1. Final design for the phone attachment

Each participant completed eight recordings, with simulated errors such as coughing, poor posture, and device mishandling, based on ATS/ERS spirometry guidelines [1].

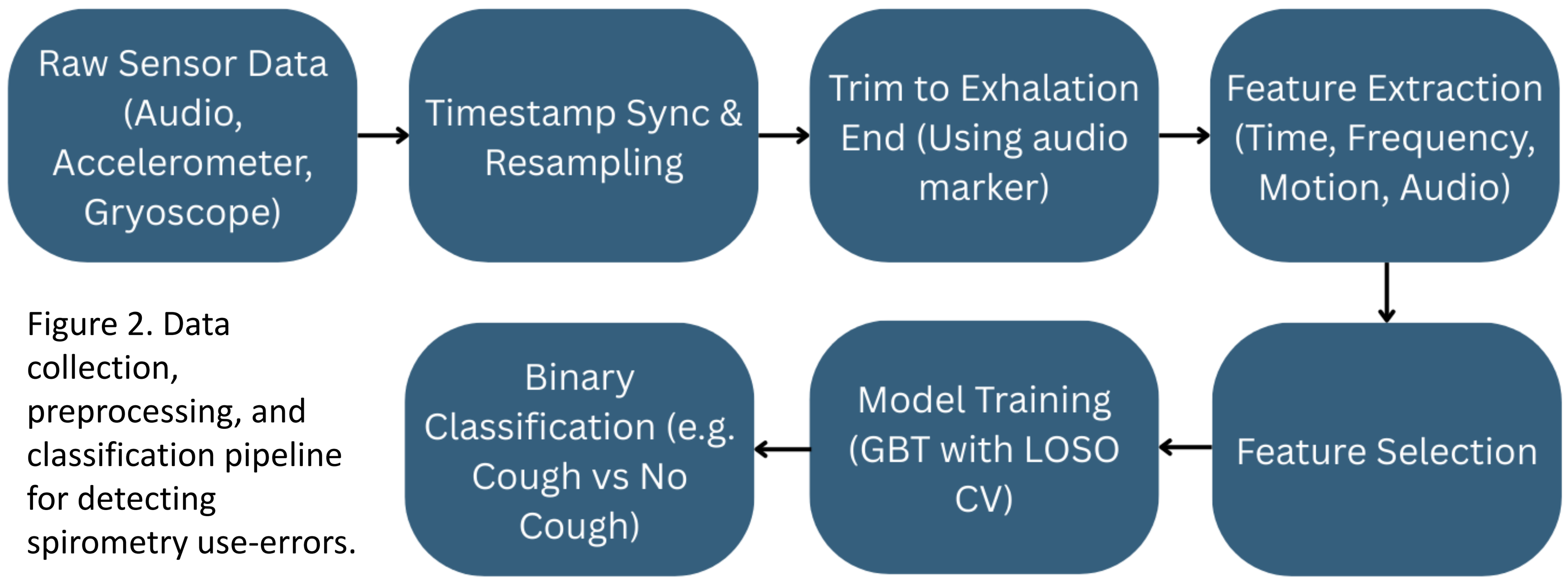

Figure 2. Data collection, preprocessing, and classification pipeline for detecting spirometry use-errors.

As shown in Figure 2, sensor data (audio, accelerometer, gyroscope) was synchronised and trimmed. Features were extracted across multiple domains—time, frequency, motion, and audio—and refined using feature selection methods. Gradient Boosted Trees (GBTs) were trained using Leave-One-Subject-Out cross-validation, as established in our previous work [2] to classify use-errors such as coughing and device mishandling. Studies have shown LOSO CV can be effective for small-sized datasets, where randomly splitting may overestimate performance through subject bias [3].

3 RESULTS

Initially, we tested windowed and non-windowed feature extraction methods, discovering that using the latter yielded better performance. Both forward and backward feature selection were outperformed by GBT-derived importance scores, which improved accuracy to 0.69 and F1-score to 0.68.

Table 1 presents the results from our binary classifiers across various error types. Overall, the models demonstrated strong performance, with Mishandling being detected most reliably (0.97 accuracy). More subtle errors, such as Bad Posture and Not Full Inhale, proved more difficult to detect, with macro F1 scores of 0.65 and 0.74, respectively. Despite these challenges, weighted metrics remained high, indicating consistent detection performance across imbalanced data. Feature importance analysis revealed that gyroscope and audio features were the most influential across all label classifications.

Table 1. The final evaluation metric results for binary classification models for each label

Label	Precision (Macro)	Recall (Macro)	F1 (Macro)	Precision (Weight-ed)	Recall (Weight-ed)	F1 (Weight-ed)
Normal	0.83	0.82	0.82	0.84	0.84	0.84
Cough	0.78	0.73	0.75	0.89	0.90	0.90
Mishandling	0.99	0.90	0.94	0.98	0.97	0.97
Tongue	0.81	0.78	0.79	0.91	0.91	0.91
Not Full Inhale	0.72	0.76	0.74	0.89	0.88	0.88
Bad Posture	0.70	0.63	0.,65	0.85	0.88	0.86

4 CONCLUSION

This study shows the potential of smartphone sensor data for detecting spirometry use-errors at home. Despite a small sample, classifiers performed well across all error types, demonstrating potential for this approach.

Future work includes improving data quality with a standardised phone mount, expanding testing to include more participants (including those with respiratory conditions), and advancing to multi-class and real-time classification.

These results highlight the promise of low-cost, multimodal smartphone systems for at-home spirometry.

REFERENCES

[1] Graham, B.L., Steenbruggen, I., Miller, M.R., Barjaktarevic, I.Z., Cooper, B.G., Hall, G.L., Hallstrand, T.S., Kaminsky, D.A., McCarthy, K., McCormack, M.C.,et al.: Standardization of spirometry 2019 update. an official american thoracic society and european respiratory society technical statement. American journal of respiratory and critical care medicine 200(8), e70–e88 (2019)

[2] Gardiner, D., Rogers, H., Lines, J., Wilson, A. Aung, M.H.: Quality Assurance for Home Spirometry using Machine Learning. In: 2025 IEEE ICTS4eHealth [Not Published Yet]

[3] Mekruksavanich, S., Hnoohom, N., Phaphan, W., Jitpattanakul, A.: A comparative study of validation methods for sensor-based human activity recognition using deep learning models. In: 2025 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON). pp. 686–691. IEEE (2025)