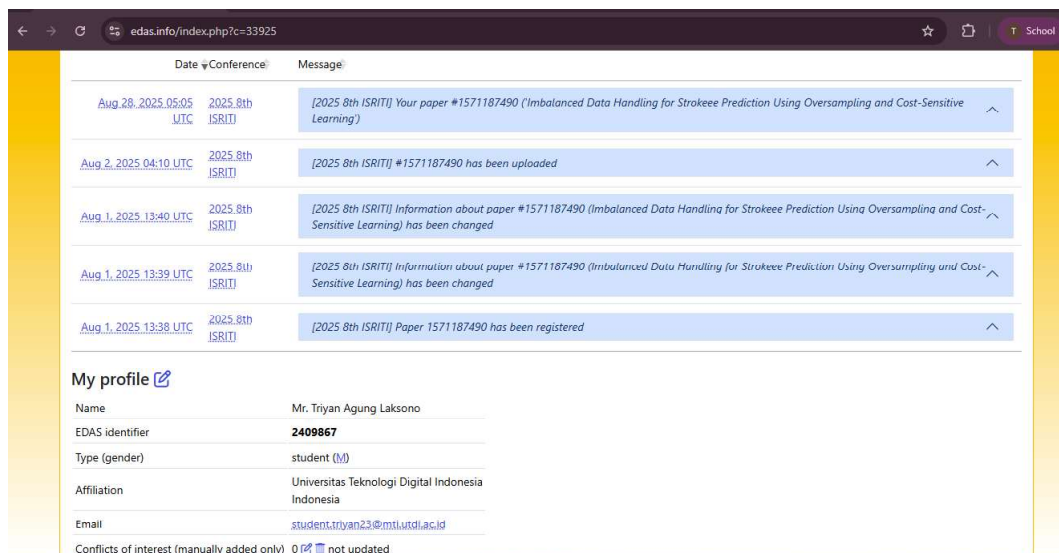


## BAB II

### PELAKSANAAN PUBLIKASI

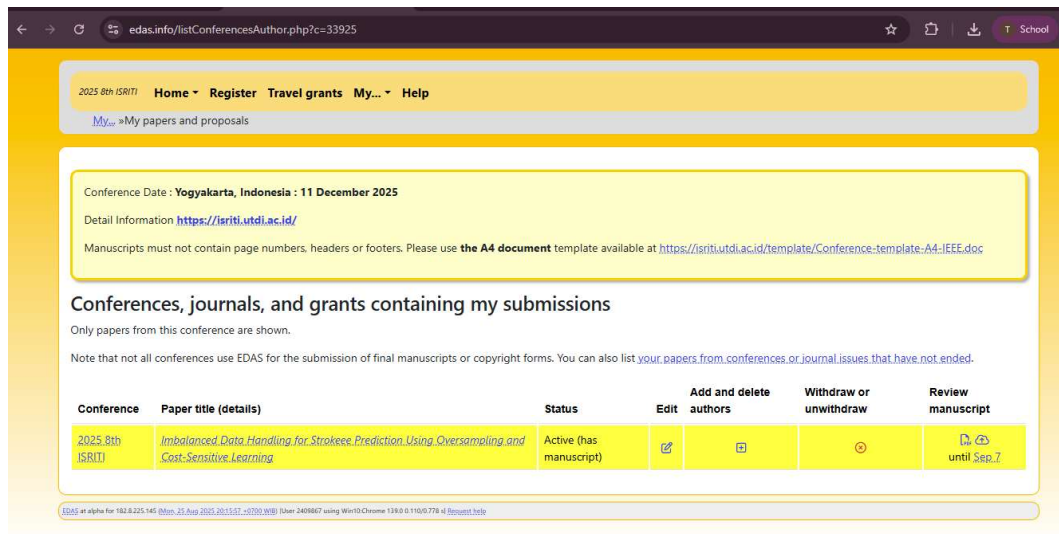
#### 2.1. Tahapan publikasi

Artikel yang saya buat dengan judul “Imbalanced Data Handling for Stroke Prediction Using Oversampling and Cost-Sensitive Learning” saya registrasikan pada “The 2025 8th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)” tanggal 1 Agustus 2025 yang dapat dilihat pada gambar 2.1 dan gambar 2.2.

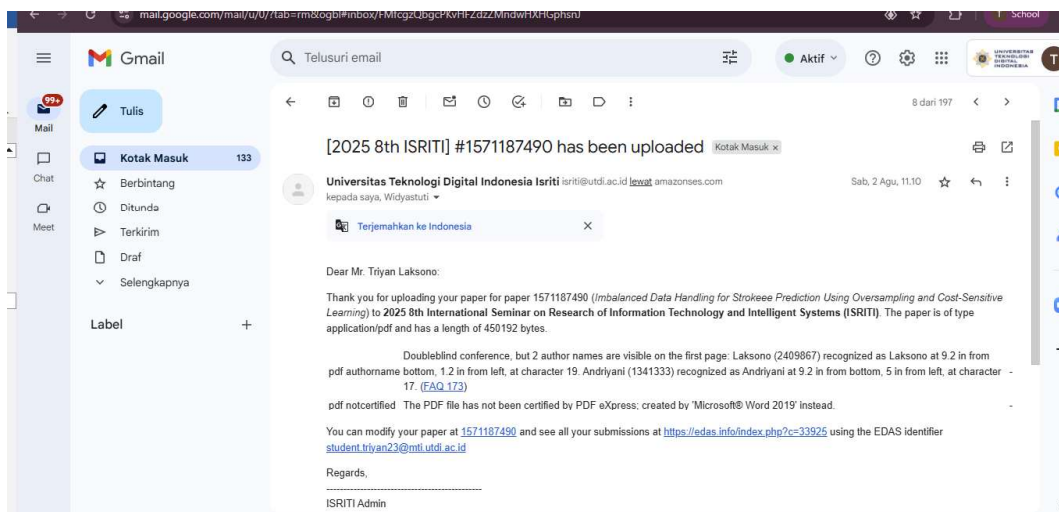


**Gambar 2. 1** Informasi Registrasi ISRITI

##### 2.1.1. Bukti submit



**Gambar 2. 2 Submit Artikel ISRTI**



**Gambar 2. 3 Notifikasi Submit ISRTI Melalui E-Mail**

### 2.1.2. Bukti review karya ilmiah dari reviewer jurnal/prosiding

#36 (1571187490): Imbalanced Data Handling for Stroke Prediction Using Oversampling and Cost-Sensitive Learning

Hide details

HbT<sub>92</sub>N

## Authors

[Triyan Agung Laksono](#) (Universitas Teknologi Digital Indonesia, Indonesia); [Widyastuti Andriyani](#) (Universitas Teknologi Digital Indonesia & UTDI, Indonesia)

45

**Paper title** *Imbalanced Data Handling for Stroke Prediction Using Oversampling and Cost-Sensitive Learning* Only the chairs can edit

Conference and track 2025 8th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) - Data and Distributed Computing

## Abstract

**Keywords** Stroke prediction; class imbalance; SMOTE; cost-sensitive learning; machine learning. Only the chairs can edit

**Topics** Soft Computing, Fuzzy Logic, and Artificial Neural Networks; Decision support systems, performance indicators, and control (Only the chairs can edit)

## Personal notes

**Roles** You are the creator and an author for this paper.  
You have authored an accepted paper in this conference.

### Status

Copyright ©

Presented by not specified 

[Review manuscript](#) [Final manuscript](#) [Presentation](#) [Additional Files, such us: similarity report, attachments, data sources, etc.](#) [Video Presentation for Back Up \(optional\)](#)

Review

Technical content and scientific rigour

Novelty and originality

Quality of presentation

Relevance and timeliness

Recommendation

Good

4 Average

3 Good

4 Excellent

5 Accept.

Detailed comments

This paper addresses the critical problem of imbalanced datasets in stroke prediction by comparing two established techniques: SMOTPE oversampling and cost-sensitive learning, using three widely adopted machine learning models (Random Forest, SVM, and XGBoost). The technical execution is sound, with a clear methodology and the valuable inclusion of SHAP-based feature importance analysis, which enhances model interpretability for medical decision-making. The results are presented in a structured way, highlighting the trade-off between improved recall for the minority class and reduced precision and overall accuracy, a realistic and important observation for clinical applications. The topic is both relevant and timely, given the increasing adoption of AI in healthcare and the persistent challenge of data imbalance in medical datasets.

To further strengthen the technical content, the paper could incorporate additional evaluation metrics such as F1-score, ROC-AUC, and precision-recall curves to provide a more complete performance picture, especially in imbalanced data contexts. Statistical significance testing, for example McNemar's test or Wilcoxon signed-rank test, would improve the robustness of the comparative analysis. In terms of novelty, the study would benefit from exploring or proposing a hybrid balancing method that combines oversampling with cost-sensitive approaches, or integrating domain-specific clinical risk factors to enrich the feature set.

On the presentation side, the paper is generally clear and well-structured; however, expanding the literature review with more recent studies from the last two years would better position this work in the current research landscape. A brief discussion on the limitations of the dataset and the potential for overfitting when using SMOTPE would add transparency and scientific rigour. Minor grammatical edits could also improve the overall readability.

Overall, the paper makes a meaningful and relevant contribution, and with the suggested improvements, it could have even stronger technical impact and presentation quality.

Editor Average

2 Editor Average

3 Editor Average

3 Editor Average

2 Likely Reject

1

Detailed comments

The term "breakout" is somewhat confusing. Is it a typo error?

reference does not conform with IEEE formatting

The limited data may somewhat affect the validity of results. This relying solely on the public Coggle stroke dataset, which may be relatively small and lack demographic diversity. Does the generalizability of the results to real-world populations? Table 1 should be discussed. How might the reduction in precision and accuracy caused by the balancing techniques affect the reliability of predictions in a clinical setting, where false positives could lead to unnecessary medical intervention or treatment?

Good

4 Average

3 Good

4 Good

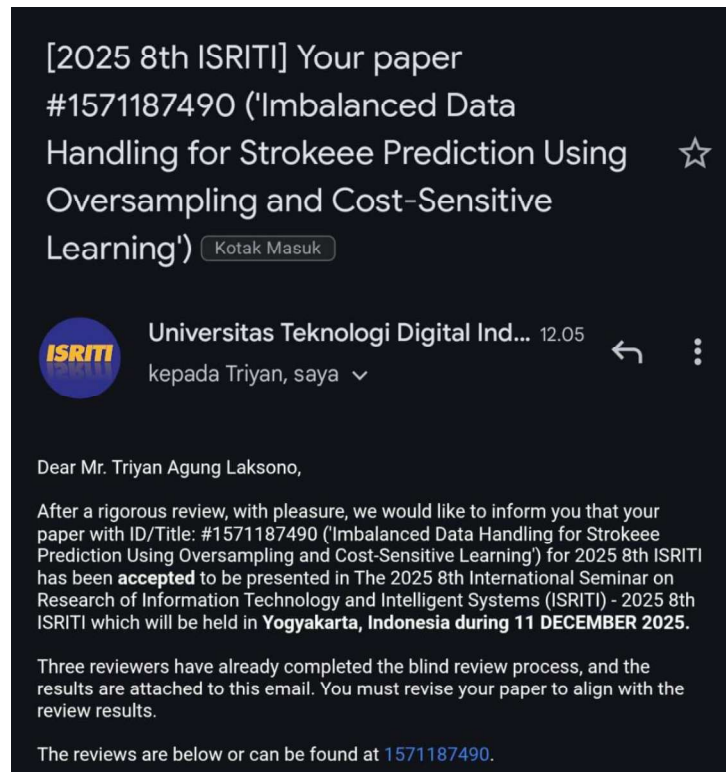
5 Positive Accept.

2

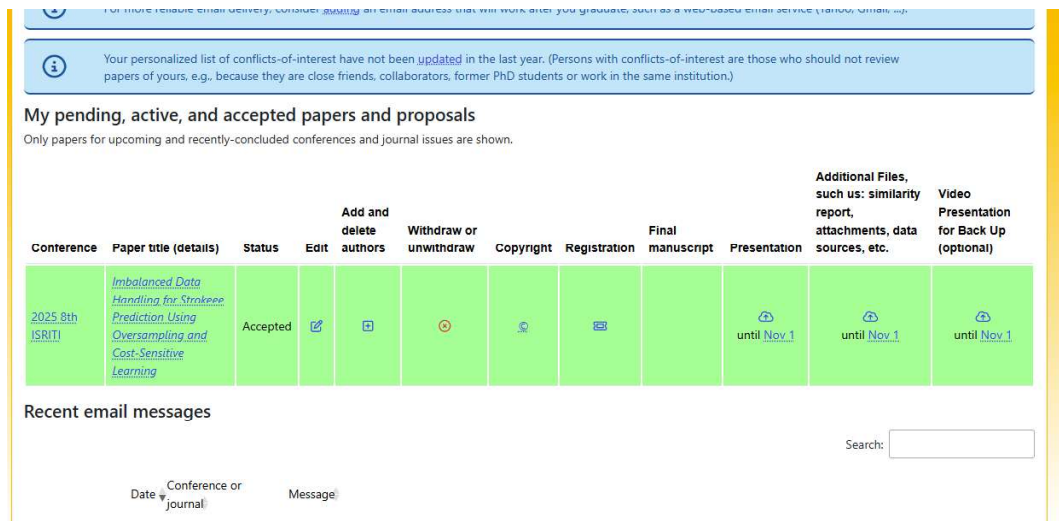
Detailed comments

**Gambar 2. 4** Halaman Review ISRITI

## 2.2. Bukti Publikasi



**Gambar 2. 5** Bukti Accepted ISRTI Melalui E-Mail



**Gambar 2. 6** Bukti Accepted Pada Aplikasi ISRTI