

**REVIEW JURNAL
PENELITIAN OBSERVASIONAL
MATA KULIAH RISET INFORMATIKA**

“Predictive Model for Extended-Spectrum β -Lactamase–Producing Bacterial Infections Using Natural Language Processing Technique and Open Data in Intensive Care Unit Environment: Retrospective Observational Study”



Andini Fitriyah Salsabilah (21081010217)

DOSEN PENGAMPU:
Assoc. Prof. Dr. Basuki Rahmat, S.Si, MT, ITS-AI

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA
TIMUR 2024**

Penelitian ini termasuk penelitian observasional karena penelitian ini menggambarkan studi retrospektif observasional yang menggunakan data sekunder dari *Medical Information Mart for Intensive Care III* (MIMIC-3), sebuah database rekam medis elektronik yang tersedia secara publik. Penelitian ini mengembangkan model prediktif untuk mendeteksi infeksi bakteri yang memproduksi *extended-spectrum β -lactamase* (ESBL) dengan memanfaatkan data yang sudah ada tanpa melakukan intervensi langsung pada pasien atau pengaturan variabel secara eksperimental.

Judul	Predictive Model for Extended-Spectrum β -Lactamase–Producing Bacterial Infections Using Natural Language Processing Technique and Open Data in Intensive Care Unit Environment: Retrospective Observational Study
Jurnal	JMIR Formative Research
Volume	8
Tahun	2024
Penulis	Genta Ito, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki
Instansi Penulis	Department of Information Science, Nara Institute of Science and Technology, Ikoma City, Japan
Abstrak	Penelitian ini mengembangkan model prediktif untuk mendeteksi infeksi bakteri yang memproduksi <i>extended-spectrum β-lactamase</i> (ESBL) di unit perawatan intensif menggunakan <i>natural language processing</i> (NLP) dan data terbuka dari MIMIC-3. Dengan memanfaatkan <i>named entity recognition</i> melalui QuickUMLS, penelitian ini menunjukkan potensi teknologi terbuka dalam membantu prediksi yang bermakna secara klinis.
Permasalahan	Tantangan yang diangkat dalam penelitian ini adalah prediksi infeksi ESBL menggunakan data terbuka dan teknologi NLP sederhana, serta keterbatasan dalam pengumpulan data medis terkait infeksi dari rekam medis elektronik.
Subjek Penelitian	Pasien ICU dalam dataset MIMIC-3 yang memiliki catatan infeksi bakteri <i>Klebsiella pneumoniae</i> atau <i>Escherichia coli</i> .
Metode Penelitian	Penelitian menggunakan metode observasional retrospektif, di mana data dari MIMIC-3 dianalisis menggunakan dua model: regresi logistik dengan regularisasi L1 dan model LightGBM. Performa model dievaluasi menggunakan ROC-AUC dan PR-AUC.
Hasil Penelitian	Dari 46.520 pasien, 3.614 di antaranya memiliki data yang lengkap. Model LightGBM memberikan hasil ROC-AUC 0,707 dan PR-AUC 0,369, menunjukkan bahwa penambahan data dari teks rekam medis dapat meningkatkan akurasi model dibandingkan hanya menggunakan data terstruktur.

Kekuatan Penelitian	Menggunakan data terbuka dari MIMIC-3, memungkinkan replikasi dan transparansi. Menerapkan pendekatan NLP untuk mengekstraksi informasi penting dari teks rekam medis yang tidak terstruktur.
Kelemahan Penelitian	Bergantung pada akurasi <i>named entity recognition</i> yang digunakan, yang mungkin menghasilkan kesalahan dalam interpretasi istilah medis. Data berasal dari satu institusi, sehingga mungkin kurang representatif untuk populasi yang lebih luas.
Kesimpulan	Studi ini berhasil membangun model prediktif yang transparan untuk infeksi bakteri ESBL menggunakan data terbuka dan teknologi NLP. Meskipun ada keterbatasan, penelitian ini menyoroti potensi penggunaan data teks medis dalam meningkatkan akurasi model prediksi klinis