

ABSTRAK

Proses pemesanan tiket kapal pada jaringan sub-agen masih bergantung pada komunikasi manual melalui aplikasi pesan instan, menciptakan inefisiensi operasional dan risiko hilangnya peluang transaksi ketika tiket yang dicari tidak tersedia. Penelitian ini merancang dan membangun sebuah aplikasi web berbasis *Proof of Concept* (POC) yang mengintegrasikan data pemesanan tiket secara *real-time* dari Platform Induk menggunakan teknik *web scraping*, serta dilengkapi dengan sistem rekomendasi berbasis *machine learning* yang menyediakan jadwal alternatif secara otomatis ketika tiket yang dicari tidak tersedia. Solusi teknis menerapkan arsitektur empat layanan terdistribusi, yaitu *Backend Service* (Flask), *Scraper Service* (Requests + BeautifulSoup4), *ML Service* (Random Forest), dan *Frontend Service* (Vue.js), yang bekerja terkoordinasi untuk menyediakan pencarian jadwal *real-time*, simulasi pemesanan berbasis *wallet*, dan rekomendasi jadwal alternatif. Model Random Forest yang dilatih pada 2.000 data sintesis yang dibangkitkan berdasarkan skema 206 data historis riil mencapai *Accuracy* 75,3%, *Precision* 80,5%, *Recall* 81,4%, *F1-Score* 80,9%, dan *AUC-ROC* 0,836, melampaui seluruh *threshold* evaluasi yang ditetapkan. Pada tahap validasi akhir, seluruh 21 skenario pengujian penerimaan dinyatakan terpenuhi (21/21 Pass) oleh dua evaluator dari pihak mitra riset, dikonfirmasi melalui Surat Konfirmasi Keterlibatan Mitra Riset. Penelitian ini membuktikan bahwa integrasi *web scraping* dan *machine learning* dalam satu platform terpadu secara teknis layak sebagai solusi otomasi pemesanan tiket kapal pada jaringan agen bertingkat.

Kata Kunci: *Web Scraping*, *Machine Learning*, Random Forest, Sistem Rekomendasi, *Proof of Concept*

ABSTRACT

The ship ticket booking process within sub-agent networks still relies on manual communication through instant messaging applications, creating operational inefficiencies and lost transaction opportunities when requested tickets are unavailable. This study designs and develops a web-based application as a Proof of Concept (POC) that integrates ticket availability data in real-time from the Main Platform using web scraping techniques, along with a machine learning-based recommendation system that automatically provides alternative schedules when requested tickets are unavailable. The technical solution employs a four-service distributed architecture — Backend Service (Flask), Scraper Service (Requests + BeautifulSoup4), ML Service (Random Forest), and Frontend Service (Vue.js) — working in coordination to deliver real-time schedule retrieval, wallet-based booking simulation, and automatic alternative schedule recommendations. The Random Forest model, trained on 2,000 synthetic records generated from the schema of 206 real historical transactions, achieved an Accuracy of 75.3%, Precision of 80.5%, Recall of 81.4%, F1-Score of 80.9%, and AUC-ROC of 0.836, surpassing all established evaluation thresholds. During the final validation phase, all 21 user acceptance testing scenarios were confirmed as fulfilled (21/21 Pass) by two evaluators from the research partner institution, evidenced by a formal Research Partner Involvement Confirmation Letter. This study demonstrates that the integration of web scraping and machine learning within a unified platform is technically feasible as an automated ticketing solution for hierarchical agent networks in the maritime transportation domain.

Keywords: Web Scraping, Machine Learning, Random Forest, Recommendation System, Proof of Concept