



Implementasi Telegram Bot berbasis OCR (Optical Character Recognition) dan NLP (Natural Language Processing) untuk Pencatatan Penjualan pada UMKM

Iman Ahsani Yasfin *, Adi Suwondo, Nur Hasanah

Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Sains Al-Qur'an, Indonesia

*Email (Penulis Korespondensi): imanahsanii@gmail.com

Abstrak. Pencatatan transaksi penjualan secara manual masih menjadi kendala bagi UMKM karena dapat menyebabkan kesalahan pencatatan, keterlambatan informasi, dan kesulitan dalam memantau kondisi usaha. Digitalisasi pencatatan diperlukan untuk mendukung pengelolaan transaksi yang lebih terstruktur dan membantu pengambilan keputusan berbasis data. Penelitian ini bertujuan merancang dan membangun sistem pencatatan penjualan berbasis cloud menggunakan Telegram Bot yang mengintegrasikan Optical Character Recognition (OCR), Natural Language Processing (NLP), fuzzy matching, Regular Expression (Regex), database, dan dashboard website pada UMKM A&W Production. Mistral OCR digunakan untuk mengekstraksi teks dari gambar catatan transaksi, sedangkan NLP berbasis fuzzy matching dan Regex digunakan untuk mengidentifikasi intent serta mengekstraksi informasi transaksi menjadi data terstruktur. Data hasil pemrosesan selanjutnya disimpan pada database Supabase dan ditampilkan melalui dashboard website untuk mendukung pemantauan penjualan. Pengujian dilakukan menggunakan Character Error Rate (CER), Word Error Rate (WER), Confusion Matrix, dan Black Box Testing. Hasil pengujian OCR menghasilkan rata-rata CER sebesar 2,65% dan WER sebesar 7,44%, sedangkan pengujian NLP memperoleh accuracy, precision, recall, dan F1-score sebesar 1,00 pada 20 data uji. Sistem juga berhasil menjalankan seluruh fungsi yang diuji sesuai kebutuhan. Integrasi teknologi tersebut mendukung digitalisasi pengelolaan UMKM dan berkontribusi pada SDG 8, SDG 9, dan SDG 12 melalui peningkatan efisiensi pencatatan, pemanfaatan teknologi, dan pengelolaan transaksi yang lebih terstruktur.

Kata kunci: Telegram Bot; Optical Character Recognition; Natural Language Processing; Fuzzy Matching, UMKM

Abstract. Manual sales transaction recording remains a challenge for MSMEs because it can lead to recording errors, delayed information, and difficulties in monitoring business conditions. Digitalization of transaction recording is needed to support more structured transaction management and data-driven decision-making. This study aims to design and develop a cloud-based sales recording system using a Telegram Bot that integrates Optical Character Recognition (OCR), Natural Language Processing (NLP), fuzzy matching, Regular Expression (Regex), a database, and a web-based dashboard for A&W Production MSME. Mistral OCR is used to extract text from transaction note images, while NLP based on fuzzy matching and Regex is used to identify intents and extract transaction information into structured data. The processed data are stored in a Supabase database and displayed through a web-based dashboard to support sales monitoring. System evaluation was conducted using Character Error Rate (CER), Word Error Rate (WER), Confusion Matrix, and Black Box Testing. The OCR evaluation achieved an average CER of 2.65% and WER of 7.44%, while the NLP evaluation achieved an accuracy, precision, recall, and F1-score of 1.00 on 20 test data. The system also successfully performed all tested

functions according to the specified requirements. The integration of these technologies supports MSME digitalization and contributes to SDG 8, SDG 9, and SDG 12 through improved recording efficiency, technology utilization, and more structured transaction management.

Keywords: Telegram Bot; Optical Character Recognition; Natural Language Processing; Fuzzy Matching, MSMEs

1. Pendahuluan

UMKM memiliki peran yang sangat penting dalam perekonomian Indonesia karena berkontribusi dalam menciptakan lapangan kerja dan meningkatkan pendapatan masyarakat (Munthe dkk., 2023). Selain itu, UMKM menyumbang lebih dari 60% terhadap Produk Domestik Bruto (PDB) nasional serta hampir 97% terhadap penyerapan tenaga kerja (Kemenko Perekonomian RI, 2025). Besarnya kontribusi tersebut menunjukkan bahwa penguatan dan digitalisasi UMKM menjadi bagian penting dalam mendukung pertumbuhan ekonomi nasional. Pemerintah juga terus mendorong UMKM untuk naik kelas melalui peningkatan kapasitas dan daya saing usaha. Dalam kondisi tersebut, pencatatan transaksi yang akurat dan terstruktur menjadi semakin penting karena data transaksi dapat digunakan sebagai dasar untuk memantau kinerja usaha, mengelola arus kas, menyusun laporan keuangan, serta mendukung pengambilan keputusan bisnis. Data keuangan yang terdokumentasi dengan baik juga dapat membantu UMKM menunjukkan kondisi usahanya ketika membutuhkan akses terhadap pembiayaan.

Meskipun memiliki peran besar dalam perekonomian, banyak UMKM masih menghadapi kendala dalam pengelolaan keuangan, terutama karena pencatatan transaksi yang masih dilakukan secara manual. Kondisi ini sering menimbulkan keterlambatan dalam penyusunan laporan keuangan, kesalahan pencatatan, dan kesulitan dalam memantau kondisi keuangan usaha (Rizky Firmansyah dkk., 2024). Akibatnya, pelaku UMKM dapat mengalami hambatan dalam mengambil keputusan bisnis yang tepat, seperti dalam penetapan harga, pengelolaan arus kas, dan perencanaan usaha ke depan (Salsabila & Febryana, 2025). Di tengah meningkatnya persaingan bisnis pada era digital, kondisi tersebut menjadi semakin mendesak karena pelaku usaha membutuhkan informasi transaksi yang cepat dan akurat untuk mengevaluasi kinerja serta menentukan keputusan berdasarkan data.

Salah satu UMKM yang mengalami kendala tersebut adalah A&W Production yang berlokasi di Desa Kacepit, Kecamatan Selopampang, Kabupaten Temanggung. A&W Production merupakan *home industry* yang bergerak di bidang produksi dan penjualan makanan ringan. Dalam menjalankan usahanya, transaksi penjualan masih dicatat secara manual menggunakan buku catatan. Cara pencatatan ini menyebabkan pengelolaan data penjualan, pemantauan pembayaran pelanggan, dan penyusunan laporan menjadi kurang efisien. Selain itu, keterlambatan penyampaian informasi kepada pemilik usaha pernah menyebabkan keterlambatan pembayaran, sedangkan proses pencarian data transaksi juga memerlukan waktu yang cukup lama dan berisiko menimbulkan kesalahan pencatatan. Romadani dkk. (2026) menjelaskan bahwa pencatatan yang tidak teratur belum dapat dimanfaatkan secara maksimal sebagai dasar penilaian kinerja dan pengambilan keputusan.

Berdasarkan permasalahan tersebut, digitalisasi pencatatan penjualan diperlukan untuk membantu UMKM mengubah catatan transaksi menjadi data yang lebih terstruktur dan mudah dipantau. Penelitian ini merancang dan mengembangkan Telegram Bot berbasis

Optical Character Recognition (OCR) dan *Natural Language Processing* (NLP) untuk membantu proses pencatatan penjualan pada UMKM A&W Production. Telegram dipilih karena menyediakan *Application Programming Interface* (API) yang memudahkan pengembangan *chatbot* serta integrasi dengan sistem tanpa perlu membangun *platform* komunikasi dari awal. Penelitian Nabila dan Abdulmajid (2026) menunjukkan bahwa Telegram dapat dimanfaatkan untuk pencatatan dan pengiriman data ke *database* secara otomatis. Sementara itu, penelitian Nicholas dkk. (2024) menunjukkan bahwa pengembangan Telegram Bot berbasis *Natural Language Processing* (NLP) mampu mendukung proses penjualan pada UMKM melalui penyediaan informasi produk, layanan pelanggan, serta respons otomatis terhadap pertanyaan pengguna.

Dalam penelitian ini, *Optical Character Recognition* (OCR) digunakan untuk mengekstraksi informasi dari dokumen gambar ke dalam bentuk data digital, sedangkan *Natural Language Processing* (NLP) digunakan untuk memahami masukan pengguna sehingga dapat diproses oleh sistem. Penelitian ini menggunakan Mistral OCR karena tersedia dalam bentuk *Application Programming Interface* (API) sehingga mudah diintegrasikan ke dalam sistem tanpa memerlukan instalasi model secara lokal (Mistral AI Team, 2025). Untuk meningkatkan akurasi pemrosesan, penelitian ini juga menerapkan *fuzzy matching* guna mengoreksi kesalahan penulisan (*typo*) dan hasil ekstraksi OCR. Penelitian Siahaan dan Muhamad Alda (2025) menjelaskan bahwa *fuzzy matching* efektif dalam menangani kesalahan penulisan, sedangkan Rahmatsyah dkk. (2025) menyatakan bahwa kombinasi OCR dan *fuzzy matching* mampu meningkatkan kualitas hasil ekstraksi teks.

Selain itu, penelitian ini menerapkan *Regular Expression* (Regex) untuk mengenali, memvalidasi, dan mengekstraksi pola teks sesuai format yang telah ditentukan. Menurut Azzahra dkk. (2025) dan Andriansyah (2025), *Regular Expression* (Regex) merupakan metode yang efektif untuk mencocokkan pola teks sehingga dapat digunakan dalam proses validasi, pencarian, dan ekstraksi data berbasis *string*. Data transaksi dari masukan pengguna kemudian diproses, disimpan ke dalam *database*, dan ditampilkan melalui *dashboard website* sehingga pemilik usaha dapat memantau penjualan dengan lebih mudah. Penelitian Salbuanda dan Esabella (2026) menjelaskan bahwa sistem pemantauan penjualan berbasis web mampu mengatasi pencatatan transaksi yang tidak terorganisir serta meningkatkan proses pemantauan dan analisis penjualan.

Meskipun penelitian sebelumnya telah menunjukkan manfaat Telegram Bot untuk pencatatan dan proses penjualan, penelitian Nabila dan Abdulmajid (2026) berfokus pada integrasi Telegram Bot dengan *Google Sheets* API, sedangkan Nicholas dkk. (2024) berfokus pada penerapan *chatbot* berbasis NLP dalam sistem penjualan. Penelitian-penelitian tersebut belum mengintegrasikan pemrosesan gambar menggunakan OCR dengan *fuzzy matching*, Regex, penyimpanan data transaksi, dan *dashboard website* dalam satu alur pencatatan penjualan. Oleh karena itu, kebaruan penelitian ini terletak pada pengembangan sistem pencatatan penjualan UMKM secara terintegrasi yang memungkinkan pengguna memasukkan transaksi melalui pesan teks maupun gambar catatan transaksi, kemudian memproses masukan tersebut menggunakan Mistral OCR, *fuzzy matching*, dan Regex, menyimpan hasilnya ke dalam *database*, serta menyajikan informasi transaksi melalui *dashboard website*. Integrasi tersebut membentuk alur pencatatan dari masukan hingga

pemantauan data dalam satu sistem yang disesuaikan dengan kebutuhan operasional UMKM A&W Production.

Pengembangan sistem ini juga memiliki keterkaitan dengan Tujuan Pembangunan Berkelanjutan (*Sustainable Development Goals*, SDGs). Penelitian ini mendukung SDG 8, khususnya target 8.3, melalui pemanfaatan teknologi digital untuk mendukung produktivitas dan pengelolaan UMKM sebagai bagian dari pertumbuhan ekonomi. Penelitian ini juga berkaitan dengan SDG 9, khususnya target 9.5, melalui penerapan teknologi OCR, NLP, *fuzzy matching*, dan integrasi sistem digital untuk meningkatkan kapasitas teknologi dan inovasi. Selain itu, digitalisasi pencatatan transaksi dan pemantauan penjualan dapat mendukung SDG 12 melalui pengelolaan kegiatan usaha yang lebih terstruktur, efisien, dan berkelanjutan. Dengan demikian, pengembangan sistem tidak hanya ditujukan untuk menyelesaikan permasalahan pencatatan pada A&W Production, tetapi juga mendukung pemanfaatan teknologi digital dalam peningkatan kapasitas dan pengelolaan UMKM.

Berdasarkan uraian tersebut, penelitian ini mengembangkan sistem pencatatan penjualan berbasis Telegram Bot yang memungkinkan pengguna mencatat transaksi melalui pesan teks maupun gambar catatan transaksi. Data yang diterima kemudian diproses secara otomatis, disimpan ke dalam *database*, dan ditampilkan pada *dashboard website* untuk memudahkan pengelolaan serta pemantauan transaksi penjualan. Telegram dipilih karena dapat digunakan melalui perangkat *mobile* sehingga lebih fleksibel bagi pengguna tanpa harus mengakses komputer. Sistem yang dikembangkan diharapkan dapat meningkatkan efisiensi pencatatan, mengurangi kesalahan *input*, serta mempermudah pemilik usaha dalam memperoleh informasi transaksi sebagai dasar pemantauan dan pengambilan keputusan.

2. Metode

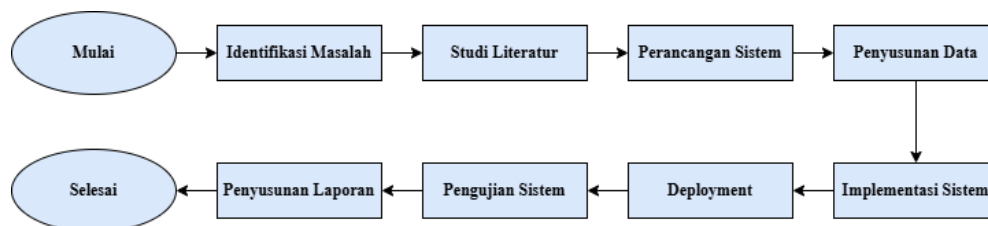
Penelitian ini merupakan penelitian terapan (*applied research*) dengan pendekatan kuantitatif eksperimental yang bertujuan mengembangkan sistem Telegram Bot berbasis *Optical Character Recognition* (OCR) dan *Natural Language Processing* (NLP) untuk mendukung pencatatan penjualan pada UMKM A&W Production. Penelitian dilaksanakan di UMKM A&W Production yang berlokasi di Desa Kacepit, Kecamatan Selopampang, Kabupaten Temanggung pada periode April hingga Juni 2026.

Dataset yang digunakan disusun secara mandiri berdasarkan kebutuhan operasional UMKM A&W Production. Dataset ini mencakup berbagai aktivitas bisnis, seperti pencatatan transaksi penjualan, pelunasan cicilan, pemantauan pembayaran, pengelolaan data pelanggan, dan penyusunan laporan penjualan. Pada bagian NLP, data terdiri atas 9 *system intent* utama, 37 *intent* spesifik (*sub-intents*), 1.019 pola atau contoh kalimat, serta 30 aturan *Regex*. Setiap pesan yang diproses akan dipetakan ke salah satu dari 9 *system intent* utama untuk menjalankan logika bisnis pada sistem. Sementara itu, 37 *intent* spesifik digunakan untuk mewakili berbagai kebutuhan pengguna, sedangkan 1.019 pola atau contoh kalimat digunakan oleh *fuzzy matcher* untuk mencocokkan masukan pengguna dengan pola yang tersedia. Selain itu, 30 aturan *Regex* digunakan untuk mengenali pola tertentu dan membantu proses ekstraksi informasi transaksi.

Penelitian ini berfokus pada perancangan dan pembangunan sistem pencatatan penjualan berbasis Telegram Bot yang mengintegrasikan OCR dan NLP. Analisis dilakukan

terhadap kinerja Mistral OCR dalam mengekstraksi informasi dari gambar transaksi serta kinerja NLP berbasis *rule-based*, *fuzzy matching*, dan *Regular Expression* (Regex) dalam mengidentifikasi *intent* dan mengekstraksi entitas.

Pengumpulan data dilakukan melalui studi literatur, wawancara, dan observasi. Studi literatur digunakan untuk memperoleh landasan teori dan referensi yang berkaitan dengan OCR, NLP, Telegram Bot, *fuzzy matching*, *Regular Expression* (Regex), serta metode evaluasi sistem. Wawancara dilakukan dengan pemilik UMKM A&W Production untuk mengidentifikasi kebutuhan sistem, sedangkan observasi dilakukan untuk memahami proses pencatatan penjualan dan karakteristik dokumen transaksi yang digunakan. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur penelitian

Alur penelitian meliputi tahap identifikasi masalah, studi literatur, pengumpulan data, *preprocessing*, *processing*, implementasi sistem, pengujian, dan penyusunan laporan. Tahap *preprocessing* meliputi *text cleaning*, *lowercase*, *tokenization*, dan normalisasi teks. Selanjutnya, tahap *processing* menerapkan *fuzzy matching* menggunakan *library* RapidFuzz dengan *threshold* sebesar 0,85 untuk menangani kesalahan penulisan pada masukan pengguna maupun hasil ekstraksi Mistral OCR. Hasil pencocokan digunakan untuk menentukan *sub-intent*, kemudian dipetakan ke salah satu *system intent* utama untuk menjalankan logika bisnis. Ekstraksi entitas dilakukan menggunakan pendekatan *rule-based* dengan *Regular Expression* (Regex) yang dirancang berdasarkan karakteristik dan format transaksi A&W Production. Sebanyak 30 pola Regex digunakan untuk mengenali informasi transaksi, seperti tanggal, pelanggan, barang, jumlah, satuan, nominal pembayaran, metode pembayaran, dan status pembayaran. Data hasil pemrosesan kemudian disusun menjadi data transaksi yang terstruktur, disimpan ke dalam *database*, dan ditampilkan pada *dashboard website*.

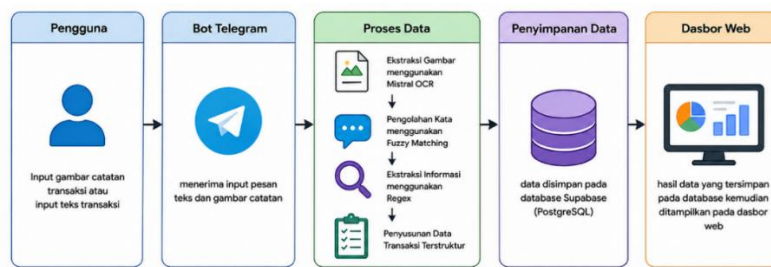
Analisis data dilakukan dengan mengevaluasi kinerja OCR dan NLP. Kinerja OCR diukur menggunakan *Character Error Rate* (CER) dan *Word Error Rate* (WER). Sementara itu, kinerja NLP dievaluasi menggunakan *Confusion Matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kemampuan sistem dalam mengidentifikasi *intent*.

3. Hasil dan Pembahasan

3.1 Implementasi Sistem Pencatatan Penjualan

Penelitian ini berhasil mengembangkan sistem pencatatan penjualan berbasis *cloud* menggunakan Telegram Bot dengan menerapkan *Optical Character Recognition* (OCR) dan *Natural Language Processing* (NLP) untuk membantu UMKM A&W Production dalam melakukan pencatatan transaksi secara digital. Sistem mengintegrasikan Telegram Bot, Mistral OCR, NLP berbasis *fuzzy matching* dan *Regular Expression* (Regex), *database* Supabase, serta *dashboard website* sehingga proses pencatatan, pengolahan, penyimpanan, dan

penyajian data transaksi dapat dilakukan secara otomatis. Arsitektur sistem yang dikembangkan ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur Sistem

3.1.2 Implementasi Telegram Bot

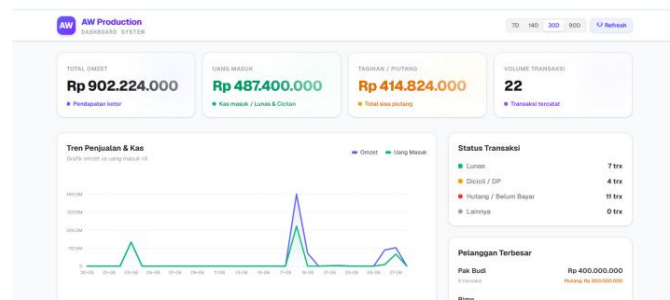
Telegram Bot berfungsi sebagai antarmuka utama untuk melakukan pencatatan transaksi melalui masukan berupa teks maupun gambar. Gambar yang dikirim pengguna diproses menggunakan Mistral OCR API untuk mengekstraksi teks, sedangkan teks hasil OCR maupun masukan langsung diproses oleh modul NLP berbasis *fuzzy matching* dan *Regular Expression* (Regex) untuk mengidentifikasi *intent* serta mengekstraksi informasi transaksi. Hasil pemrosesan kemudian ditampilkan sebagai konfirmasi kepada pengguna sebelum disimpan ke *database* dan ditampilkan pada *dashboard website*. Implementasi Telegram Bot ditunjukkan pada Gambar 3.



Gambar 3. Implementasi Telegram Bot

3.1.3 Implementasi Dashboard Website

Dashboard website dikembangkan sebagai media pemantauan transaksi penjualan yang terhubung langsung dengan *database*. Informasi yang ditampilkan meliputi total omzet, jumlah transaksi, total pembayaran yang diterima, nilai piutang pelanggan, daftar transaksi terbaru, serta grafik penjualan berdasarkan periode tertentu. Dengan demikian, pemilik usaha dapat memantau kondisi penjualan secara lebih cepat dan akurat. Tampilan *dashboard website* ditunjukkan pada Gambar 4.



Gambar 4. *Dashboard Website*

3.1.4 Hasil Pengujian Optical Character Recognition (OCR)

Pengujian OCR dilakukan untuk mengevaluasi kemampuan Mistral OCR dalam mengekstraksi informasi dari gambar catatan transaksi penjualan. Pengujian menggunakan delapan gambar tulisan tangan yang memuat isi transaksi yang sama, tetapi memiliki karakteristik penulisan yang berbeda. Hasil ekstraksi kemudian dibandingkan dengan *ground truth* menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER). Hasil pengujian OCR ditunjukkan pada Tabel 1.

Tabel 1. Hasil Pengujian OCR

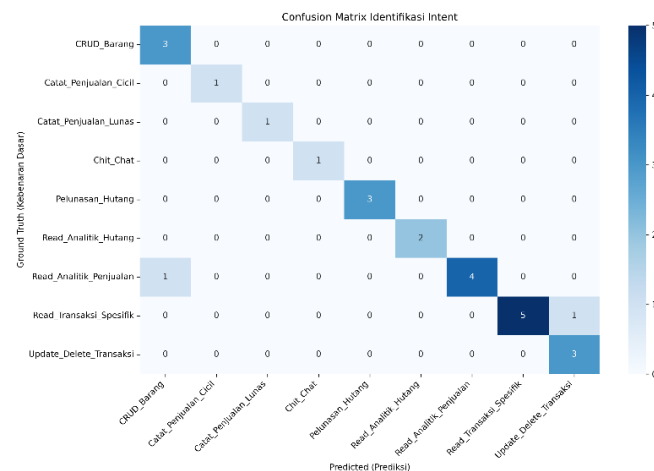
No.	Nama Gambar	CER	WER
1	Gambar 1	1.48%	4.76%
2	Gambar 2	0.00%	0.00%
3	Gambar 3	1.48%	7.14%
4	Gambar 4	9.36%	19.05%
5	Gambar 5	2.96%	14.29%
6	Gambar 6	3.94%	7.14%
7	Gambar 7	1.48%	4.76%
8	Gambar 8	0.49%	2.38%

Berdasarkan hasil pengujian, Mistral OCR memperoleh rata-rata CER sebesar 2,65% dan WER sebesar 7,44%. Nilai tersebut menunjukkan bahwa hasil ekstraksi memiliki tingkat kesalahan yang relatif rendah pada delapan sampel yang diuji. Namun, jumlah sampel yang digunakan masih terbatas sehingga hasil tersebut belum sepenuhnya merepresentasikan variasi tulisan tangan, kualitas gambar, tingkat kejelasan tulisan, dan kondisi pencatatan yang lebih beragam. Oleh karena itu, hasil pengujian OCR dalam penelitian ini perlu dipahami sebagai evaluasi awal terhadap kemampuan Mistral OCR pada karakteristik dokumen transaksi A&W Production dan belum dapat digeneralisasikan untuk seluruh kondisi tulisan tangan pengguna.

3.1.5 Hasil Pengujian Identifikasi *Intent*

Pengujian identifikasi *intent* dilakukan untuk mengevaluasi kemampuan sistem dalam mengenali maksud pengguna berdasarkan masukan yang diberikan melalui Telegram Bot. Pengujian menggunakan 20 data uji yang terdiri atas berbagai jenis transaksi dan perintah sistem, seperti pencatatan transaksi, pelunasan hutang, pencarian transaksi, analisis penjualan, dan pengelolaan data. Evaluasi dilakukan menggunakan *Confusion Matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan

bahwa sistem mampu mengidentifikasi *intent* dengan tingkat akurasi yang tinggi sehingga setiap masukan pengguna dapat diproses sesuai fungsi yang diinginkan. *Confusion Matrix* hasil pengujian ditunjukkan pada Gambar 5.



Gambar 5. *Confusion Matrix* Identifikasi Intent

Berdasarkan Gambar 5, sistem berhasil mengidentifikasi seluruh data uji dengan benar sehingga memperoleh nilai *accuracy* sebesar 0,92. Selain itu, hasil evaluasi *precision*, *recall*, dan *F1-score* untuk setiap kelas ditunjukkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Identifikasi Intent

Nama Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
CRUD Barang	0,75	1,00	0,86
Catat Penjualan Cicil	1,00	1,00	1,00
Catat Penjualan Lunas	1,00	1,00	1,00
Chit Chat	1,00	1,00	1,00
Pelunasan Hutang	1,00	1,00	1,00
Read Analitik Hutang	1,00	1,00	1,00
Read Analitik Penjualan	1,00	0,80	0,89
Read Transaksi Spesifik	1,00	0,83	0,91
Update Delete Transaksi	0,75	1,00	0,86

3.1.6 Hasil Pengujian *Black Box*

Pengujian *Black Box* dilakukan untuk memastikan seluruh fungsi sistem berjalan sesuai dengan kebutuhan yang telah ditetapkan. Pengujian meliputi proses pencatatan transaksi menggunakan teks dan gambar, konfirmasi hasil ekstraksi, pencarian transaksi, pembaruan data, penghapusan data, pengelolaan data master, serta *dashboard website*. Hasil pengujian menunjukkan bahwa seluruh fungsi menghasilkan keluaran sesuai dengan yang diharapkan sehingga sistem dapat digunakan dengan baik dalam mendukung proses pencatatan transaksi penjualan. Ringkasan hasil pengujian ditunjukkan pada Tabel 3.

Tabel 3. Hasil Pengujian *Black Box*

No.	Fitur yang Diuji	Hasil Pengujian	Status
1	<i>Input</i> transaksi teks	Sistem berhasil mengenali <i>intent</i> dan entitas transaksi serta menampilkan ringkasan transaksi.	Berhasil
2	<i>Input</i> transaksi gambar	Sistem berhasil melakukan OCR pada gambar yang valid dan memberikan notifikasi pada gambar yang tidak dapat diproses.	Berhasil
3	Konfirmasi penyimpanan	Data berhasil disimpan atau dibatalkan sesuai pilihan pengguna.	Berhasil
4	Perubahan, pembaruan, dan penghapusan transaksi	Sistem berhasil mengubah, memperbarui, dan menghapus data transaksi.	Berhasil
5	Pencarian dan riwayat transaksi	Sistem berhasil menampilkan data transaksi sesuai pencarian dan riwayat transaksi.	Berhasil
6	<i>Dashboard website</i>	Dashboard berhasil menampilkan data transaksi, omzet, uang masuk, dan tagihan.	Berhasil
7	Pengelolaan data master	Sistem berhasil menambah, mengubah, dan menghapus data barang, satuan, dan metode pembayaran.	Berhasil
8	Validasi <i>input</i>	Sistem berhasil mendeteksi data yang belum lengkap dan memberikan notifikasi kepada pengguna.	Berhasil
9	Cek hutang	Sistem berhasil menampilkan daftar hutang dan informasi hutang pelanggan sesuai permintaan.	Berhasil

3.2 Pembahasan

Hasil penelitian menunjukkan bahwa sistem pencatatan penjualan berbasis *cloud* berhasil diimplementasikan dengan mengintegrasikan Telegram Bot, Mistral OCR, modul NLP, *database* Supabase, dan *dashboard website*. Integrasi tersebut memungkinkan proses pencatatan transaksi dilakukan secara otomatis, mulai dari penerimaan masukan pengguna, ekstraksi informasi dari teks maupun gambar, identifikasi *intent*, penyimpanan data ke *database*, hingga penyajian informasi pada *dashboard website*. Dengan demikian, sistem dapat mendukung efisiensi proses pencatatan dan mempermudah pemantauan transaksi pada UMKM A&W Production dibandingkan metode manual yang sebelumnya digunakan.

Hasil pengujian OCR memperoleh nilai rata-rata *Character Error Rate* (CER) sebesar 2,65% dan *Word Error Rate* (WER) sebesar 7,44%. Nilai tersebut menunjukkan bahwa Mistral OCR mampu mengekstraksi teks dari catatan transaksi tulisan tangan dengan tingkat kesalahan yang rendah pada data uji yang digunakan. Temuan ini sejalan dengan penelitian Sukarna dkk. (2025) yang menyatakan bahwa semakin rendah nilai CER dan WER, semakin baik kemampuan OCR dalam mengenali teks.

Pengujian identifikasi *intent* memperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 0,92. Hasil tersebut menunjukkan bahwa metode NLP mampu mengidentifikasi *intent* dengan baik pada data uji yang digunakan, meskipun masih terdapat sebagian data yang belum diklasifikasikan sesuai dengan kelas sebenarnya. Kombinasi *fuzzy matching* dan *Regular Expression* (Regex) digunakan untuk mendukung proses identifikasi tersebut. *Fuzzy matching* berperan dalam menangani variasi dan kesalahan penulisan pada masukan pengguna maupun hasil ekstraksi OCR, sedangkan Regex digunakan untuk mengenali pola transaksi dan mengekstraksi informasi sesuai dengan format yang telah ditentukan. Temuan ini sejalan dengan penelitian Siahaan dan Muhamad Alda (2025), Rahmatsyah dkk. (2025), Azzahra dkk. (2025), dan Andriansyah (2025) yang menunjukkan bahwa *fuzzy matching* dan Regex dapat digunakan untuk mendukung proses pengolahan dan pencocokan teks.

Hasil pengujian tersebut menunjukkan bahwa integrasi OCR dan NLP dapat mendukung proses pencatatan transaksi dari masukan berupa teks maupun gambar. Namun, kinerja sistem masih dipengaruhi oleh karakteristik data yang digunakan. Pendekatan NLP yang diterapkan masih berbasis aturan (*rule-based*), sehingga kemampuan sistem dalam menangani kalimat yang lebih kompleks dan variasi masukan di luar pola yang telah ditentukan masih terbatas.

Kesimpulan

Penelitian ini berhasil mengembangkan sistem pencatatan penjualan berbasis *cloud* menggunakan Telegram Bot yang mengintegrasikan OCR, NLP, *database*, dan *dashboard website* pada UMKM A&W Production. Sistem mampu menerima input teks dan gambar, mengekstraksi informasi menggunakan Mistral OCR, mengidentifikasi *intent* dengan *fuzzy matching* dan Regex, serta menyimpan dan menampilkan data transaksi. *Black Box Testing* menunjukkan seluruh fungsi berjalan sesuai kebutuhan. Pengujian OCR menghasilkan CER sebesar 2,65% dan WER sebesar 7,44%, sedangkan pengujian identifikasi *intent* memperoleh *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 0,92. Sistem mendukung digitalisasi pencatatan dan pemantauan transaksi serta relevan dengan SDG 8.3, 9.5, dan 12.2. Keterbatasan penelitian terletak pada penggunaan delapan sampel OCR dan jumlah data uji NLP yang terbatas. Penelitian selanjutnya disarankan menggunakan dataset yang lebih besar dan beragam serta mengembangkan NLP yang mampu menangani variasi masukan lebih kompleks.

Ucapan Terima Kasih

Penulis menyampaikan apresiasi kepada UMKM A&W Production, Temanggung, atas kesempatan, kerja sama, dan penyediaan data yang mendukung pelaksanaan penelitian ini. Penulis juga mengucapkan terima kasih kepada seluruh pihak yang telah memberikan

dukungan teknis dan kontribusi selama proses pengembangan sistem, pengujian, hingga penyelesaian artikel ini.

Daftar Pustaka

- Andriansyah. (2025). *Validasi ekspresi matematika dengan Regex dan Stack untuk memeriksa keseimbangan tanda kurung*. A. Colin.
- Azzahra, F., Muhyiddin, Hayat, A. M., & Anggraeni, D. (2025). Penerapan Natural Language Processing dan Regular Expressions dalam validasi artikel. *Journal of Muhammadiyah's Application Technology*, 4(3), 335–342. <https://doi.org/10.26618/dkpw7933>
- Kementerian Koordinator Bidang Perekonomian Republik Indonesia. (2025, Januari 30). *Pemerintah dorong UMKM naik kelas, tingkatkan kontribusi terhadap ekspor Indonesia*. <https://www.ekon.go.id/publikasi/detail/6152/pemerintah-dorong-umkm-naik-kelas-tingkatkan-kontribusi-terhadap-ekspor-indonesia>
- Mistral AI Team. (2025, Maret 6). *Mistral OCR*. <https://mistral.ai/news/mistral-ocr/>
- Munthe, A., Yarham, M., Siregar, R., dkk. (2023). Peranan usaha mikro kecil menengah terhadap perekonomian Indonesia. *Jurnal Mutiara Ilmu Akuntansi*, 2(3).
- Nabila, K., & Abdulmajid, N. W. (2026). Enhancing field activity reporting through a real-time Telegram chatbot integrated with the Google Sheets API. *Integrated Science Education Journal*, 7(2), 233–243. <https://doi.org/10.37251/isej.v7i2.2682>
- Nicholas, A., Duha, Y., dkk. (2024). Penerapan *chatbot* berbasis NLP pada sistem penjualan alat elektronik dan bangunan. *Jurnal Mahasiswa Aplikasi Teknologi Komputer dan Informasi*, 6(3), 167–175.
- Muhammad Zaky Rahmatsyah, Anik Nur Handayani, & Jevri Tri Ardiansah. (2025). Comparative analysis of OCR methods integrated with fuzzy matching for food ingredient detection in Japanese packaged products. *Indonesian Journal of Data and Science*, 6(2), 301–312. <https://doi.org/10.56705/ijodas.v6i2.257>
- Rizky Firmansyah, M., Cristina Santoso, A., Farah, A., dkk. (2024). Pengaruh pencatatan akuntansi manual dengan pencatatan digital di era globalisasi dalam suatu usaha Snack Rehan Demangan Bangkalan. *Jurnal Multidisiplin Akuntansi*, 2(7). <https://doi.org/10.62281>
- Rizma Reyhana Putri, A., & Dwi Indriyanti, A. (2023). Evaluasi usability user interface dan user experience pada aplikasi M.Tix dengan metode Usability Testing (UT) dan System Usability Scale (SUS).
- Romadani, Z. W., Astuti, S. Y., & Febrianti, D. (2026). Pendampingan pembuatan pencatatan keuangan untuk meningkatkan kinerja UMKM Takao Takoyaki.
- Salbuanda, S., & Esabella, S. (2026). Rancang bangun sistem monitoring dan evaluasi penjualan pada outlet UMKM Ayam Geprek Berkah berbasis web. *Prosiding Seminar Nasional Inovasi dan Riset Multidisiplin*, 1(1).
- Salsabila, A., & Febryana, J. (2025). Pengaruh sistem informasi akuntansi terhadap kualitas laporan keuangan UMKM di era digitalisasi. *Journal of UKMC National Seminar on Accounting Proceeding*, 4(1).
- Siahaan, N. B., & Muhamad Alda. (2025). Implementasi *rule-based chatbot* dan *fuzzy matching* dalam sistem informasi layanan haji dan umrah di PT Alfata Wafiqah Wisata. *Rabit:*

Jurnal Teknologi dan Sistem Informasi Univrab, 10(2), 991-1008.
<https://doi.org/10.36341/rabit.v10i2.6413>

Sukarna, R. H., Safira, F., Marzulyanti, H. D., dkk. (2025). Studi komparatif model OCR berbasis AI untuk dokumen cetak dan tulisan tangan.

Valerian, F. R., Syarief, M., Abdul Fatah, D., dkk. (2025). Klasifikasi tingkat obesitas menggunakan metode GBM dan Confusion Matrix. *Jurnal Mahasiswa Teknik Informatika*, 9(2).

Wiwin Sry Adinda Banjarnahor, & S. (2025). *Buku ajar: Praktik pengujian perangkat lunak*. https://www.google.co.id/books/edition/BUKU_AJAR_PRAKTIK_PENGUJIAN_PERANGKAT_LU/UrCcEQAAQBAJ

ZiFlow Team. (2026, April 19). *Telegram bot vs WhatsApp bot for business in 2026: Which should you pick?* <https://ziflow.me/en/blog/telegram-bot-vs-whatsapp-bot-2026>

CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International).

This license allows users to share and adapt an article, even commercially, as long as appropriate credit is given and the distribution of derivative works is under the same license as the original. That is, this license lets others copy, distribute, modify and reproduce the Article, provided the original source and Authors are credited under the same license as the original.

