AI-Powered Counterfeit Detection and Risk Summarization

Capstone Project untuk Student Development Initiative (IBM & Hacktiv8)

Oleh Lathif Ramadhan

Table of contents

01

Latar Belakang Masalah

Tantangan Produk Palsu di E-commerce

1. Dampak Merusak Produk Palsu:
Produk palsu menimbulkan kerugian finansial yang besar bagi produsen asli dan merusak reputasi brand. Lebih penting lagi, produk palsu seringkali tidak memenuhi standar keamanan dan kualitas, membahayakan kesehatan dan keselamatan konsumen (misalnya, kosmetik palsu, suku cadang otomotif yang tidak standar, obat palsu).

2. Skala dan Kompleksitas di E-commerce:
Platform e-commerce modern memiliki jutaan listingan produk yang diperbarui setiap hari. Skala ini membuat pengawasan manual menjadi sangat sulit, memakan waktu, dan tidak efisien. Penjual palsu juga terus beradaptasi dengan metode deteksi tradisional.

3. Pentingnya Kepercayaan dan Integritas Platform:
Keberadaan produk palsu secara masif merusak kepercayaan konsumen terhadap platform e-commerce. Konsumen yang kecewa atau merasa tertipu mungkin beralih ke platform lain. Menjaga integritas platform adalah kunci untuk keberlanjutan bisnis e-commerce.

4. Krusialnya Deteksi Proaktif:
Mengingat dampak negatif produk palsu dan tantangan pengawasan manual, deteksi produk palsu secara proaktif dan otomatis menjadi sangat krusial. Sistem deteksi yang efektif dapat membantu platform mengidentifikasi dan menghapus listingan berisiko dengan cepat, melindungi konsumen, dan menjaga reputasi.

02

Tujuan Proyek

Klasifikasi Akurat & Penjelasan Risiko Cerdas

- **Tujuan utama proyek ini** adalah mengembangkan sistem cerdas yang dapat membantu platform e-commerce dalam memerangi peredaran produk palsu. Sistem ini memiliki dua pilar utama:
  - **Tujuan 1: Klasifikasi Produk Palsu**. Membangun model Machine Learning yang efektif dan akurat untuk secara otomatis mengidentifikasi listingan produk yang berpotensi palsu berdasarkan data transaksional dan atribut produk.
  - **Tujuan 2: Ringkasan Risiko Cerdas (Menggunakan IBM Granite).** Mengintegrasikan model bahasa besar (LLM) IBM Granite untuk mengubah data terstruktur dari produk yang terdeteksi palsu menjadi narasi ringkas dan mudah dipahami yang menjelaskan faktor-faktor risiko utamanya. Ini bertujuan untuk memberikan konteks yang kaya dan mempercepat proses investigasi dan pengambilan keputusan oleh tim Trust & Safety.

- **Melalui kombinasi klasifikasi prediktif dan penjelasan berbasis AI**, proyek ini bertujuan untuk meningkatkan efisiensi operasional, meminimalkan kerugian, dan yang terpenting, membangun serta menjaga kepercayaan konsumen terhadap platform.

03

Metodologi Proyek

Proyek ini mengikuti 4 alur kerja (pipeline) data science yang sistematis, terbagi menjadi fase-fase berikut:

Fase 1: Pembersihan Data & Feature Engineering

- Melakukan standardisasi data kategorikal (nama brand)
- Menangani outlier (IQR capping)
- Mengkonversi tanggal (listing_date) dan membuat fitur turunan (listing_age_days)
- Menciptakan fitur baru (purchase_to_view_ratio, is_price_below_avg)
- Serta memeriksa missing values dan duplikat.

Fase 2: Analisis Data Eksploratif (EDA)

- Menganalisis distribusi kelas target
- Memvisualisasikan hubungan fitur-fitur kunci dengan status keaslian (harga, rating, asal pengiriman, dll.)
- Mengeksplorasi korelasi numerik
- Menganalisis hubungan fitur kategorikal/boolean dengan target

Fase 3: Pemodelan & Perbandingan Model (Classification)

- Mempersiapkan data (split train/test, preprocessing: StandardScaler, OneHotEncoder)
- Membangun dan melatih 6 model klasifikasi (Logistic Regression, KNN, SVM, Random Forest, XGBoost, LightGBM)
- Mengevaluasi performa dengan metrik komprehensif (Accuracy, Precision, Recall, F1, AUC, dll.)
- Melakukan Cross-Validation
- Tuning hyperparameter (GridSearchCV pada Random Forest)
- Menganalisis Feature Importance

Fase 4: Integrasi IBM Granite (Summarization Cerdas)

- Mengintegrasikan IBM Granite via Replicate/Langchain
- Membuat prompt template kontekstual
- Membuat fungsi query
- Mendemonstrasikan penggunaan LLM untuk menghasilkan ringkasan naratif risiko dari data terstruktur produk palsu.

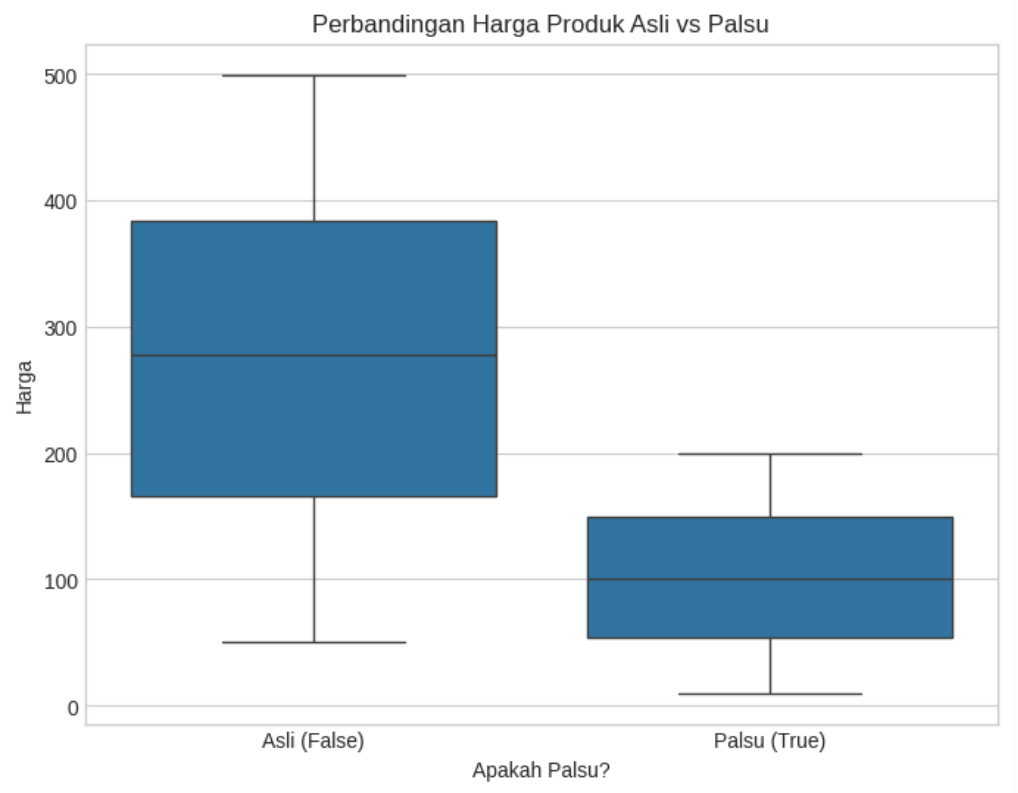
04

Temuan Kunci dari EDA

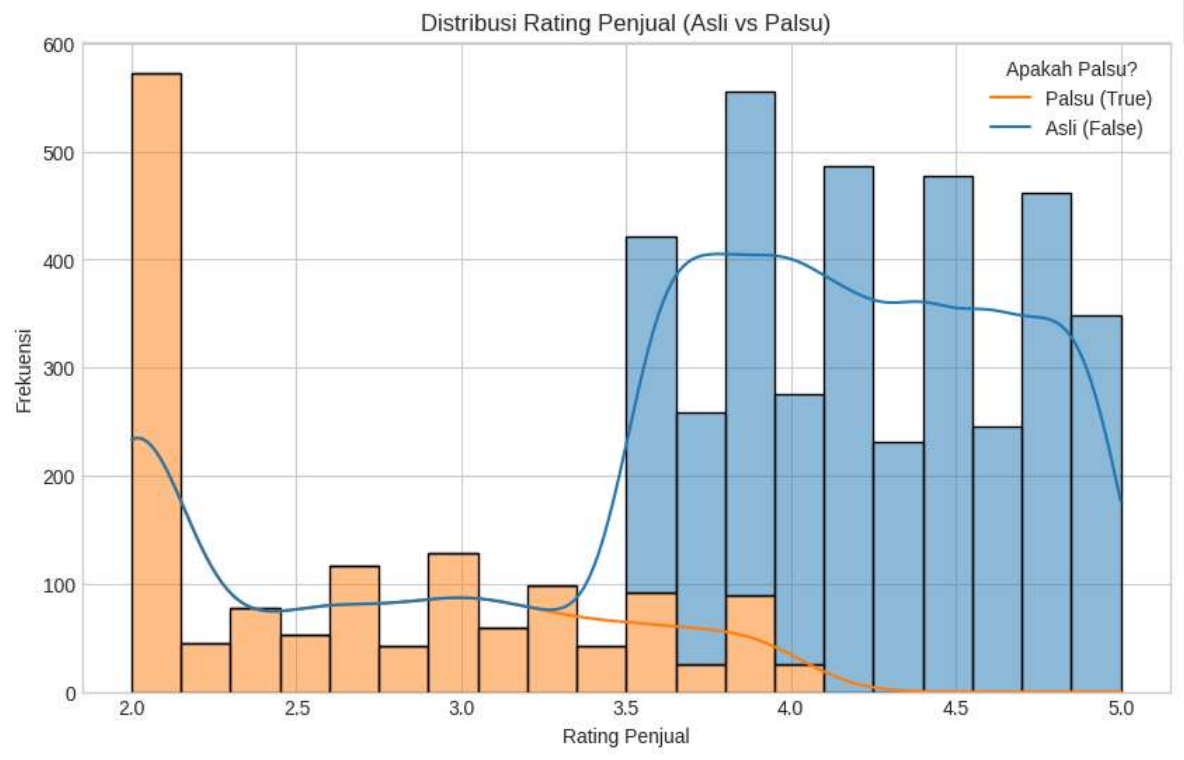
Insight Awal dari Analisis Data

Analisis data eksploratif (EDA) mengungkap beberapa pola penting yang membedakan produk asli dan palsu:

1. **Harga:**
Boxplot menunjukkan bahwa produk palsu cenderung memiliki distribusi harga yang lebihrendah dibandingkan produk asli. Fitur is_price_below_avg yang kita buat mengkonfirmasi bahwa produk yang harganya jauh di bawah rata-rata kategori memiliki kemungkinan tinggi untuk palsu.

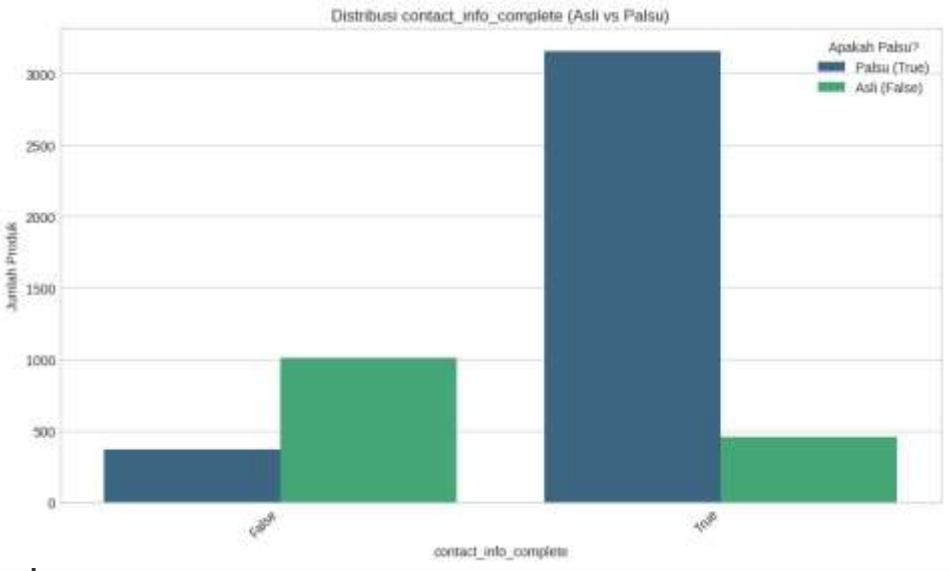


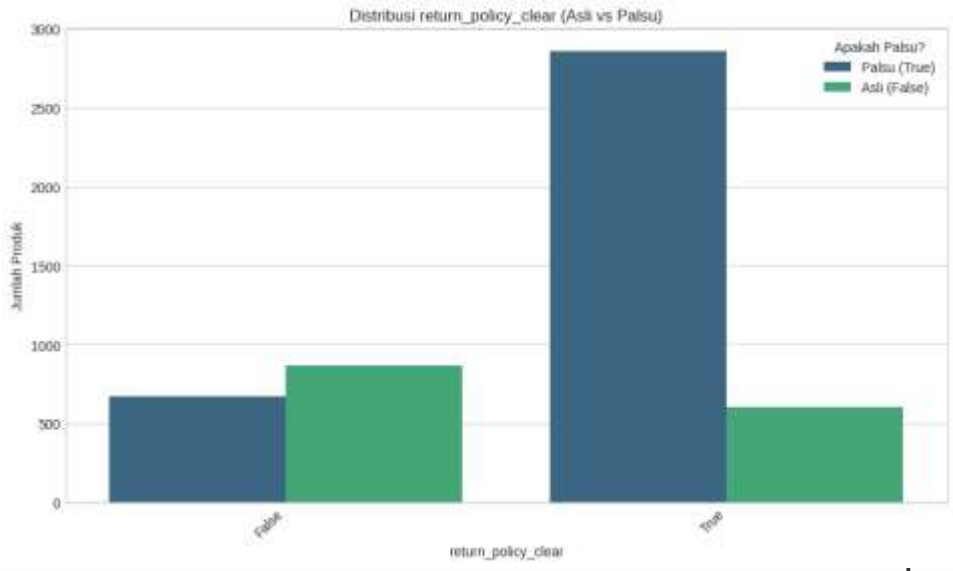
2. Rating & Ulasan Penjual: Histogram dan statistik deskriptif menunjukkan bahwa penjual produk palsu umumnya memiliki rating (seller_rating) yang lebih rendah dan jumlah ulasan (seller_reviews) yang jauh lebih sedikit dibandingkan penjual asli.



3. Informasi Penjual & Kebijakan:

Analisis fitur boolean (contact_info_complete, return_policy_clear) menunjukkan korelasi kuat; produk palsu sangat sering terkait dengan penjual yang informasinya tidak lengkap atau kebijakan pengembaliannya tidak jelas.





4. Asal Pengiriman:
Meskipun tidak seekstrim fitur boolean, beberapa asal pengiriman (shipping_origin) menunjukkan proporsi produk palsu yang lebih tinggi dibandingkan yang lain.

5. Fitur Lainnya:
Fitur seperti product_images, description_length, shipping_time_days, spelling_errors, dan domain_age_days juga menunjukkan perbedaan distribusi yang signifikan antara produk asli dan palsu, mengindikasikan nilainya sebagai prediktor.

Temuan-temuan ini memberikan dasar yang kuat untuk pemilihan fitur dan ekspektasi performa model klasifikasi. Visualisasi seperti boxplot harga vs status, histogram rating penjual, dan analisis proporsi fitur boolean adalah contoh grafik yang bisa Anda tampilkan di slide ini.

05

Hasil Pemodelan Klasifikasi

Perbandingan Performa Model Klasifikasi

Kami melatih dan mengevaluasi enam model klasifikasi populer untuk memprediksi status keaslian produk:

- Logistic Regression
- K-Nearest Neighbors (KNN)
- Support Vector Machine (SVM)
- Random Forest
- XGBoost
- LightGBM

Hasil yang Luar Biasa:
Pada dataset ini, semua model yang diuji mencapai performa yang sempurna pada data uji:

- Akurasi (Accuracy): 100%
- Presisi (Precision): 100%
- Recall: 100%
- F1-Score: 100%
- AUC ROC: 100%
- Balanced Accuracy: 100%
- Precision-Recall AUC: 100%
- Confusion Matrix menunjukkan 0 False Positif dan 0 False Negatif.

Performa sempurna ini, yang juga dikonfirmasi oleh Cross-Validation, menunjukkan bahwa fitur-fitur dalam dataset ini sangat kuat dan diskriminatif, memungkinkan pemisahan kelas yang efektif.

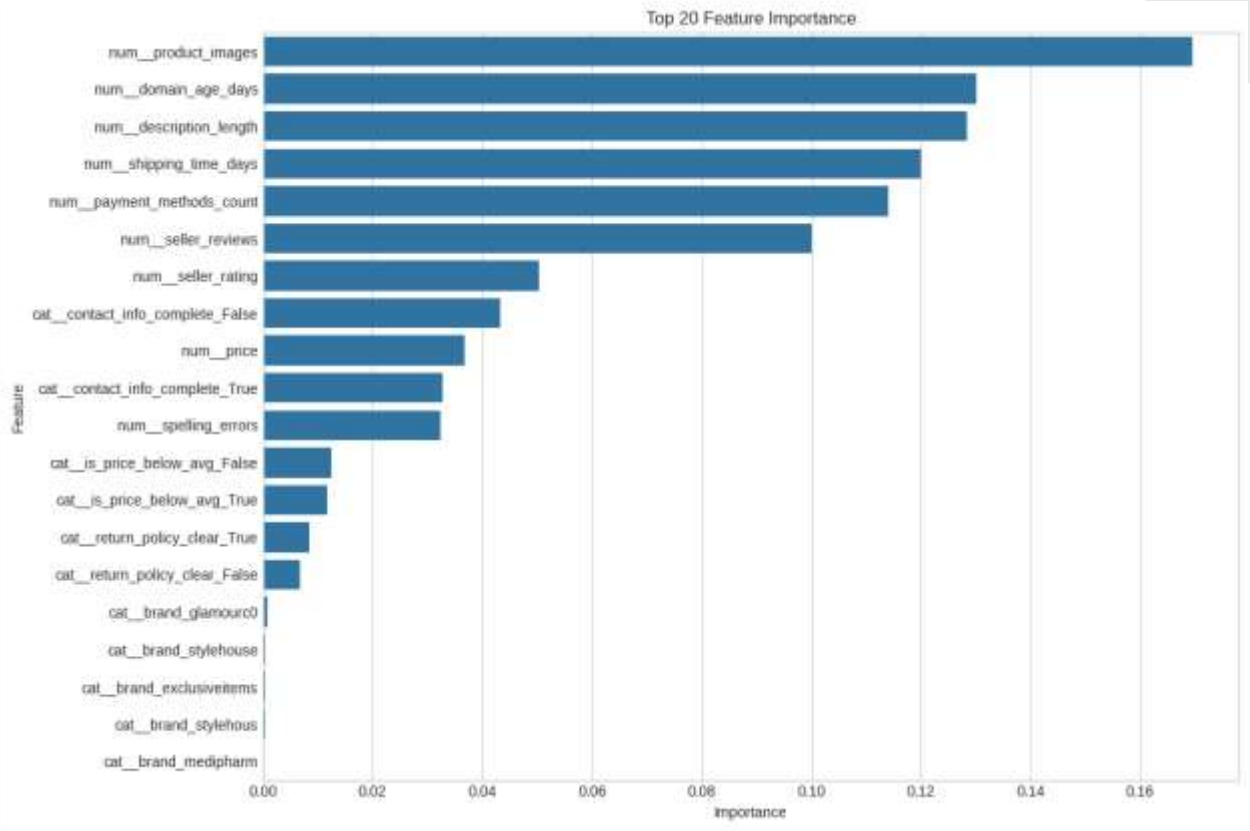
Model Pilihan: Meskipun semua model menunjukkan performa sempurna, model Random Forest (setelah tuning hyperparameter) dipilih sebagai model "juara" untuk integrasi lebih lanjut karena kemampuannya menyediakan analisis Feature Importance yang informatif.

06

Feature Importance

Fitur Paling Berpengaruh dalam Deteksi Palsu

Analisis Feature Importance dari model Random Forest (atau model berbasis tree lainnya yang menunjukkan performa tinggi) mengidentifikasi fitur-fitur kunci yang paling berpengaruh dalam memprediksi apakah suatu produk adalah palsu atau tidak.



Fitur-fitur teratas meliputi:

- **`product_images`**: Jumlah gambar produk yang sedikit (terutama 1 atau 2) adalah indikator kuat.
- **`domain_age_days`**: Umur domain atau akun penjual yang masih muda (baru dibuat) sangat relevan.
- **`description_length`**: Deskripsi yang sangat pendek cenderung terkait dengan produk palsu.
- **`shipping_time_days`**: Waktu pengiriman yang lama merupakan sinyal risiko.
- payment_methods_count: Jumlah metode pembayaran yang terbatas.
- **`seller_reviews & seller_rating`**: Jumlah ulasan yang sedikit dan rating yang rendah dari penjual.
- **`contact_info_complete`**: Penjual palsu seringkali tidak memiliki informasi kontak lengkap.
- **`price & is_price_below_avg`**: Harga yang jauh di bawah rata-rata kategori atau harga produk itu sendiri.
- **`spelling_errors`**: Keberadaan kesalahan ejaan dalam deskripsi.

Fitur-fitur ini secara logis relevan dengan karakteristik penjual dan listingan produk palsu. Mereka memberikan sinyal kuat kepada model untuk membedakan antara produk asli dan palsu, menjelaskan mengapa model klasifikasi dapat mencapai performa yang sangat tinggi pada dataset ini.

07

DEMO: IBM Granite in Action! (Summarization Cerdas)

IBM Granite: Menjelaskan Risiko dari Data Terstruktur

Di bagian ini, kami mendemonstrasikan penggunaan model bahasa besar (LLM) IBM Granite dalam cara yang inovatif. Alih-alih hanya meringkas teks naratif yang panjang, kami menggunakannya untuk tugas yang lebih kompleks: menghasilkan penjelasan dari data terstruktur.

Input: LLM menerima data terstruktur dari produk yang terdeteksi palsu oleh model klasifikasi. Data ini berupa nilai-nilai spesifik dari fitur-fitur kunci (misalnya, harga: $15, rating penjual: 2.1, waktu pengiriman: 35 hari, contact_info_complete: False, is_price_below_avg: True, dll.).

Proses: Berdasarkan konteks proyek yang diberikan dalam prompt dan data terstruktur yang diterimanya, IBM Granite melakukan reasoning untuk mengidentifikasi pola dan faktor risiko utama yang ada dalam data tersebut.

Output: LLM menghasilkan ringkasan naratif dalam bahasa alami yang menjelaskan mengapa produk tersebut dicurigai sebagai produk palsu, menyoroti faktor-faktor risiko utama. Contoh output:

> "Produk ini dicurigai palsu karena harganya sangat rendah ($15), rating penjualnya buruk (2.1), waktu pengiriman sangat lama (35 hari), dan informasi kontak penjual tidak lengkap." (Contoh Ringkasan)

Nilai Tambah: Penggunaan IBM Granite untuk tugas ini sangat berharga:

- Mempercepat Investigasi: Tim Trust & Safety dapat dengan cepat memahami alasan di balik peringatan produk palsu tanpa harus menganalisis semua data mentah.
- Meningkatkan Efisiensi: Mengotomatisasi pembuatan laporan risiko yang biasanya membutuhkan waktu manual.
- Komunikasi Efektif: Menjelaskan temuan model ML yang kompleks dalam format yang mudah dipahami oleh pengguna non-teknis.

08

Kesimpulan & Rekomendasi

Kesimpulan dan Langkah Selanjutnya

Kesimpulan:

- Proyek ini berhasil membangun model klasifikasi yang sangat akurat (mencapai 100% akurasi, presisi, recall, dan F1-score pada data uji) untuk deteksi produk palsu pada dataset yang digunakan.
- Analisis Feature Importance mengidentifikasi fitur-fitur kunci yang sangat diskriminatif, seperti jumlah gambar produk, usia domain penjual, panjang deskripsi, waktu pengiriman, rating/ulasan penjual, dan kelengkapan informasi kontak, sebagai prediktor utama produk palsu.
- Integrasi dengan IBM Granite berhasil mendemonstrasikan potensi LLM untuk menghasilkan ringkasan naratif risiko dari data terstruktur, yang sangat berharga untuk mempercepat proses investigasi manual.

Rekomendasi (Konkret & Actionable):

- **Implementasi Model Klasifikasi**: Mengitegrasikan model klasifikasi terbaik (Random Forest atau model lain dengan performa sempurna) ke dalam pipeline deteksi real-time di platform e-commerce untuk menandai listingan berisiko secara otomatis.
- **Integrasi IBM Granite**: Memanfaatkan API IBM Granite untuk secara otomatis menghasilkan ringkasan risiko naratif bagi tim Trust & Safety setiap kali model mendeteksi produk palsu. Ini akan mengurangi beban kerja manual dan mempercepat pengambilan keputusan.
- **Validasi Berkelanjutan**: Melakukan validasi model secara berkala menggunakan data terbaru dari lingkungan produksi untuk memastikan performa tetap optimal seiring waktu dan tren pemalsuan yang berkembang.
- **Eksplorasi Lebih Lanjut**: Mempertimbangkan untuk menguji model pada dataset deteksi penipuan yang lebih besar, lebih kompleks, atau yang mencakup jenis penipuan lain untuk menguji generalisasi model.
- **Pengembangan Antarmuka**: Membangun antarmuka sederhana (misalnya, dashboard internal) yang menampilkan listingan yang ditandai sebagai palsu beserta ringkasan risiko yang dihasilkan oleh IBM Granite untuk memudahkan tim operasional.

Thanks!

For more information:

- **Email**: datasciencelatieg@gmail.com
- **LinkedIn:** https://www.linkedin.com/in/lathiframadhan/
- **Github Repository Link:** *https://github.com/LatiefDataVisionary/counterfeit-detection-capstone*
- **Raw Dataset Link:** https://www.kaggle.com/datasets/aimlveera/counterfeit-product-detection-dataset/data
- **Notebook Link:** https://github.com/LatiefDataVisionary/counterfeit-detection-capstone/blob/main/notebooks/notebook.ipynb