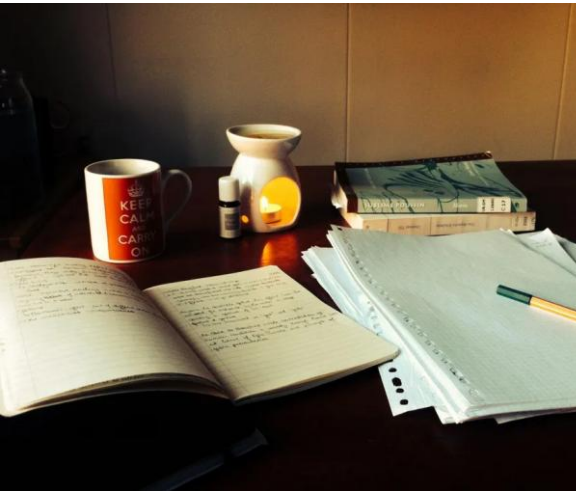


Identifikasi Marka Linguistik Pembeda Teks Generatif AI dan Teks Manusia Menggunakan *SHAP* (*SHapley Additive exPlanations*)

Muhammad Hidayat Nurwahid – 22081010300

Agenda

Pendahuluan
Research Gap
Mindmap
Metodologi Penelitian
Penutup



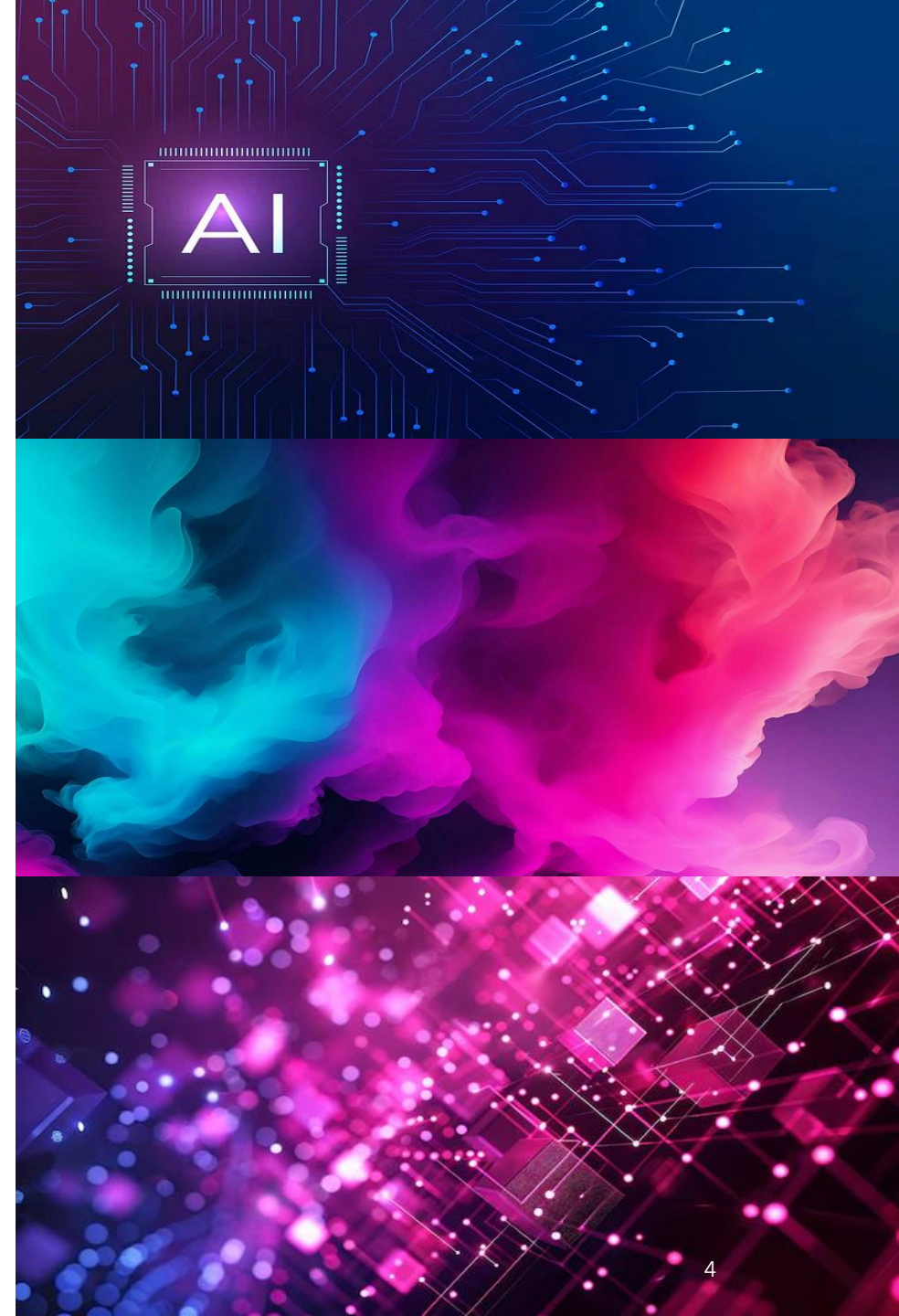
Pendahuluan

Latar Belakang Masalah dan Rumusan Masalah



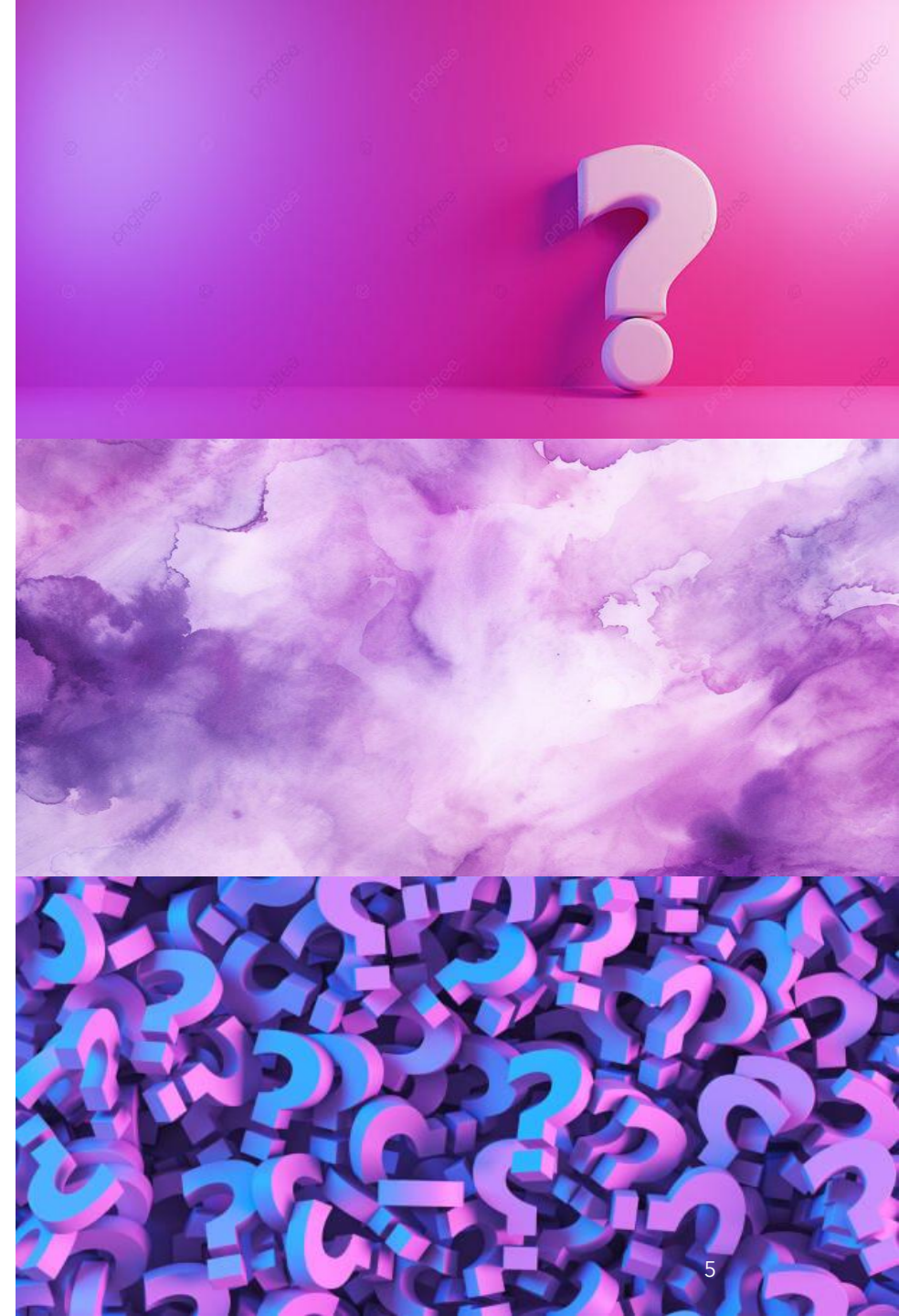
Latar Belakang Masalah

- Kehadiran AI Generatif (GAI) seperti GPT dan Gemini secara fundamental mengubah cara konten dibuat, menghasilkan teks yang sangat mirip tulisan manusia (Saha, 2025; Eapen dkk., 2023).
- Integritas karya ilmiah terancam oleh praktik plagiarisme yang semakin canggih dan sulit dilacak (Eke, 2023).
- Mayoritas alat deteksi AI yang ada beroperasi sebagai "kotak hitam" (*black box*).
- Masalah utamanya bukan ketiadaan detektor, melainkan **ketiadaan penjelasan**.



Rumusan Masalah

1. Marka linguistik apa saja (leksikal, sintaksis) yang menjadi pembeda paling signifikan antara teks buatan GAI dan tulisan manusia?
2. Bagaimana metode SHAP (*SHapley Additive exPlanations*) dapat digunakan untuk mengidentifikasi dan memvisualisasikan kontribusi setiap marka linguistik terhadap keputusan model?
3. Sejauh mana konsistensi marka linguistik tersebut dapat diandalkan pada genre teks yang berbeda?



Research Gap

Detektor yang ada diketahui memiliki "akurasi lintas-domain yang buruk". Belum ada penelitian yang secara sistematis menguji konsistensi marka linguistik pembeda di berbagai genre teks yang berbeda.

Automatic Detection of Machine Generated Text: A Critical Survey (2020)

Penelitian yang ada terlalu berfokus pada akurasi performa dan beroperasi sebagai "kotak hitam", tanpa menjelaskan mengapa sebuah model berhasil mendeteksi teks AI.

Detecting AI Generated Text Based on NLP and Machine Learning Approaches (2024)

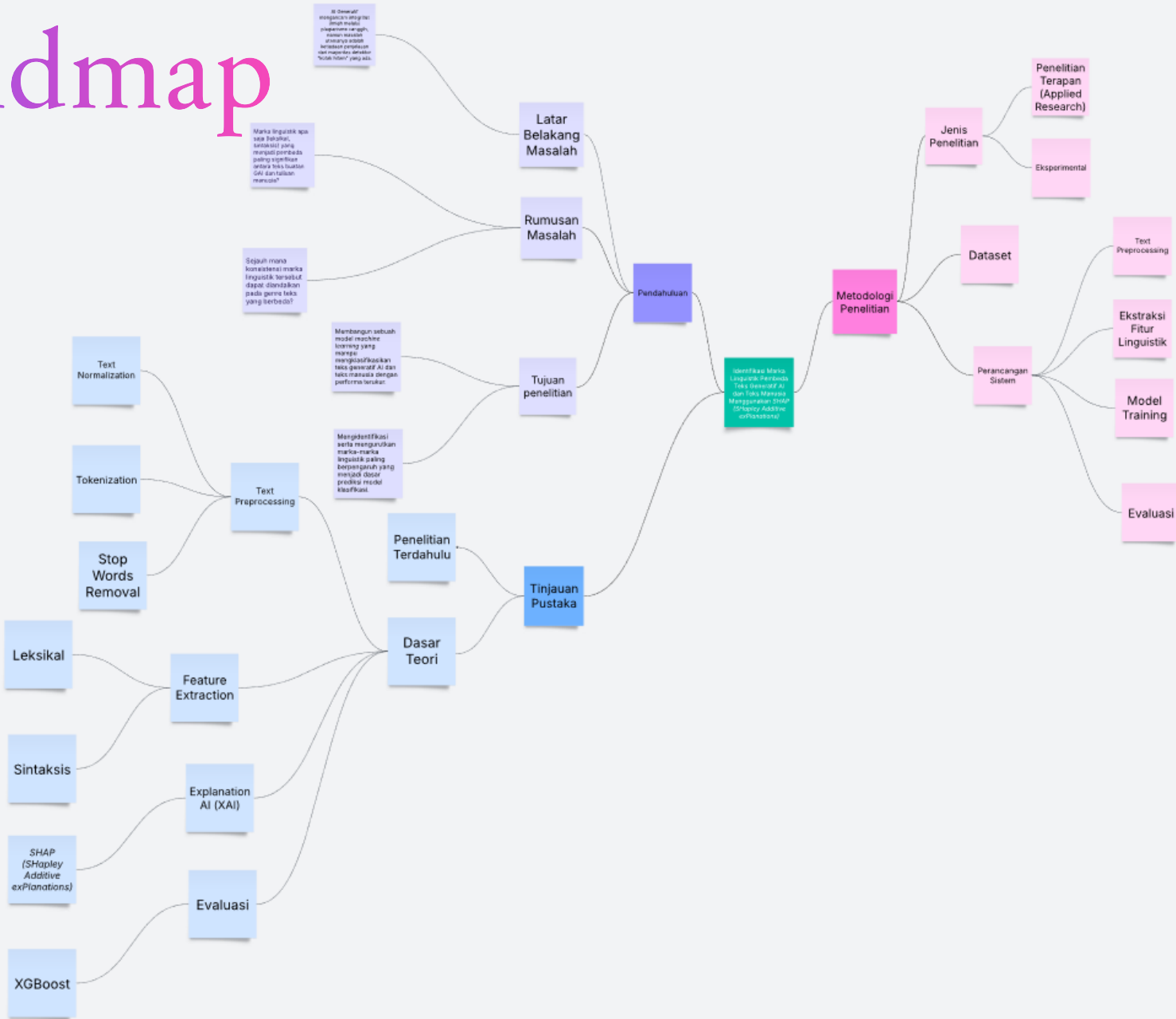
Detecting ChatGPT: A Survey of the State of Detecting ChatGPT-Generated Text (2023)

Analisis pembeda teks AI dan manusia yang ada masih bersifat kualitatif dan umum (misalnya, AI "lebih formal" atau "keragaman leksikal rendah"). Belum ada penelitian yang mengukur kontribusi spesifik dari tiap marka linguistik (leksikal, sintaksis, stilistik) terhadap keputusan model.

AI-Generated Text Detection and Source Identification (2025)

Kebutuhan akan detektor yang dapat ditafsirkan (*interpretable*) telah secara eksplisit diserukan oleh penelitian sebelumnya.

Mindmap



Metodologi Penelitian

Jenis Penelitian

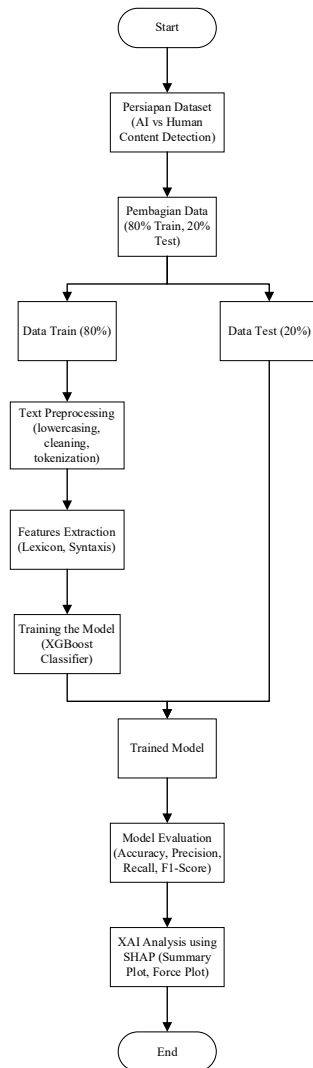
Penelitian ini merupakan penelitian terapan (*applied research*) dengan desain eksperimental kuantitatif. Dalam konteks ini, masalah praktis tersebut adalah kegagalan alat deteksi AI konvensional yang beroperasi sebagai "kotak hitam" (*black box*), sehingga tidak memiliki transparansi dan akuntabilitas.

Dataset

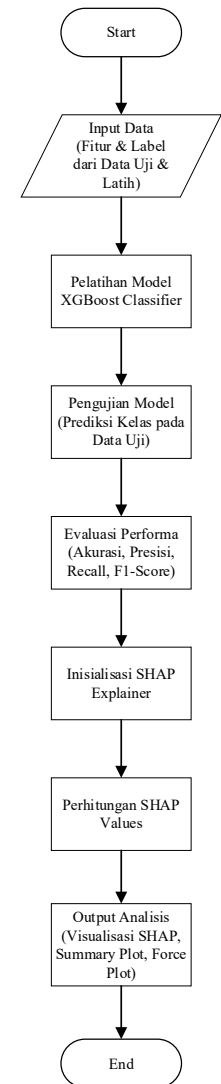
Penelitian ini menggunakan data sekunder yang diperoleh dari Kaggle. Dataset yang digunakan adalah dataset "AI vs Human Content Detection" yang dikurasi oleh Pratyush Puri.

Metodologi Penelitian

Perancangan Sistem:



Evaluasi dan Analisis:





Thank you