

# BAB 1 PENDAHULUAN

## 1.1. Latar Belakang

Dalam rangka menangani permasalahan yang terkait dengan prediksi sektor industri perusahaan berdasarkan data laporan keuangan, akan dibahas masalah pengelompokan perusahaan ke dalam sektor industri yang tepat berdasarkan data laporan keuangan. Tugas akhir ini mengambil pendekatan dengan menggunakan metode *Linear Discriminant Analysis* (LDA). Di Bursa Efek Indonesia, terdapat 12 sektor industri yang telah diklasifikasikan oleh *Indonesia Stock Exchange Industrial Classification* (IDX-IC).

Sejumlah penelitian yang relevan dilakukan sebelum penelitian ini dilakukan oleh Hans Van Der Heijden. Pada tahun 2022, Hans Van Der Heijden mengevaluasi proyeksi kinerja saham perusahaan di sektor industri Amerika Utara (NAICS) dengan menggunakan data laporan keuangan masing-masing entitas. Dalam penelitiannya, studi ini menerapkan metode *Linear Discriminant Analysis* (LDA) dan metode *Random Forest* untuk membandingkan pendekatan non-linear dan linear dalam prediksi. Temuannya mengindikasikan bahwa metode *Random Forest*, sebagai pendekatan non-linear, menunjukkan tingkat akurasi yang lebih tinggi dibandingkan dengan metode LDA. Dengan demikian, dapat disimpulkan bahwa *Random Forest* lebih efektif dalam memprediksi sektor industri dibandingkan dengan LDA [1].

Selain *Random Forest*, metode XGBoost juga merupakan salah satu pendekatan non-linear yang sering digunakan dalam klasifikasi dan prediksi berbasis laporan keuangan. XGBoost merupakan algoritma berbasis pohon keputusan yang dioptimalkan untuk meningkatkan akurasi dan efisiensi perhitungan dalam permasalahan klasifikasi. Beberapa penelitian menunjukkan bahwa XGBoost memiliki performa yang baik dalam menangani data keuangan. Studi oleh Chen dan Guestrin (2016) menunjukkan bahwa XGBoost mampu memberikan hasil klasifikasi yang lebih akurat dibandingkan dengan metode lain dalam berbagai kasus, termasuk prediksi keuangan [2]. Sementara itu, penelitian

lain oleh Li et al. (2021) mengungkapkan bahwa XGBoost lebih unggul dibandingkan model tradisional dalam memprediksi kebangkrutan perusahaan berdasarkan laporan keuangan [3].

Sejumlah penelitian juga telah dilakukan untuk mengevaluasi penerapan metode klasifikasi berbasis laporan keuangan, salah satunya adalah penelitian mengenai penggunaan Multiple Discriminant Analysis (MDA) dalam memprediksi financial distress perusahaan. MDA, yang merupakan bentuk umum dari Linear Discriminant Analysis (LDA), telah digunakan dalam analisis perusahaan sektor industri barang konsumsi yang terdaftar di Bursa Efek Indonesia. Hasil penelitian menunjukkan bahwa metode ini mampu membedakan perusahaan yang mengalami kesulitan keuangan dan yang tidak, sehingga dapat menjadi pendekatan yang efektif dalam analisis laporan keuangan untuk prediksi kinerja perusahaan [4].

Dalam tugas akhir ini, penelitian akan memfokuskan diri pada penggunaan algoritma LDA untuk mengatasi masalah prediksi sektor industri perusahaan. Dengan memanfaatkan fitur-fitur dalam laporan keuangan sebagai prediktor, algoritma ini diharapkan dapat memberikan hasil prediksi yang akurat dan dapat diinterpretasikan. Selain itu, untuk membandingkan efektivitas pendekatan linear dan non-linear, penelitian ini juga akan menambahkan metode XGBoost sebagai pembanding. Dengan adanya perbandingan ini, diharapkan dapat diperoleh pemahaman yang lebih dalam mengenai efektivitas kedua algoritma dalam menyelesaikan permasalahan klasifikasi sektor industri berdasarkan data laporan keuangan.

LDA dipilih dalam penelitian ini karena berfokus pada pendekatan linear, mengingat belum adanya penelitian sebelumnya yang secara spesifik mengangkat topik ini berdasarkan laporan keuangan. Pendekatan linear dalam klasifikasi sektor industri masih jarang dieksplorasi dalam literatur, sehingga penelitian ini diharapkan dapat memberikan kontribusi baru dalam memahami efektivitas metode LDA dalam konteks tersebut. Selain itu, LDA tetap memiliki keunggulan dalam interpretabilitas dan efisiensi komputasi, menjadikannya pilihan yang relevan dalam analisis laporan keuangan.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan yang disajikan pada bagian sebelumnya berikut merupakan rumusan masalah yang disajikan sebagai berikut :

1. Bagaimana implementasi metode *Linear Discriminant Analysis* (LDA) dan *Extreme Gradient Boosting* (XGBoost) dalam memprediksi sektor industri perusahaan di Bursa Efek Indonesia berdasarkan data laporan keuangan?
2. Bagaimana perbandingan kinerja metode *Linear Discriminant Analysis* (LDA) dan XGBoost dalam memprediksi sektor industri perusahaan di Bursa Efek Indonesia berdasarkan data laporan keuangan?
3. Bagaimana implementasi metode *Permutation Importance* dalam mengidentifikasi fitur-fitur apa saja yang berpengaruh dalam laporan keuangan perusahaan yang dapat memberikan kontribusi signifikan dalam meningkatkan keberhasilan prediksi sektor industri menggunakan algoritma LDA dan XGBoost?.

## 1.3. Tujuan dan Manfaat

Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan algoritma *Linear Discriminant Analysis* (LDA) sebagai alat analitik yang efektif dalam memprediksi sektor industri perusahaan berdasarkan data laporan keuangan dan juga membandingkan metode LDA dengan XGBoost sebagai metode pendekatan non-linearnya. Selain itu, penelitian ini juga berfokus pada analisis mendalam untuk mengidentifikasi fitur-fitur yang memberikan kontribusi signifikan terhadap peningkatan akurasi prediksi, sehingga diharapkan dapat memberikan wawasan yang lebih komprehensif mengenai pengaruh variabel keuangan terhadap klasifikasi sektor industri perusahaan.

**Tabel 1.1. Tabel keterkaitan antara tujuan, pengujian dan kesimpulan.**

no.	Tujuan	Pengujian	Kesimpulan
1	Mengimplementasikan algoritma Linear Discriminant Analysis (LDA) dan XGBoost untuk memprediksi sektor industri perusahaan berdasarkan laporannya.	Pengujian Model LDA dan XGBoost dengan mengkombinasikan fitur historis dan fundamental dilakukan menggunakan metrik akurasi dan F1 Score	Model prediksi LDA dan XGBoost berhasil dikembangkan

2	Menganalisis perbandingan dari dua pendekatan algoritma yang berbeda yaitu metode algoritma LDA dan XGBoost dalam memprediksi sektor industri perusahaan berdasarkan laporan keuangan	Menganalisis perbandingan dari hasil implementasi pengujian model LDA dan XGBoost	Hasil akurasi model yang paling tinggi dianggap model yang berhasil untuk memprediksi sektor industri saham perusahaan berdasarkan laporan keuangan
3	Mengidentifikasi fitur-fitur yang memiliki dampak paling signifikan dalam meningkatkan akurasi prediksi sektor industri perusahaan menggunakan algoritma LDA dan XGBoost	Pengujian Model LDA dan XGBoost dengan memanfaatkan metode <i>Permutation Importance</i> untuk melihat fitur - fitur yang paling berpengaruh dalam memprediksi sektor industri	Fitur fitur yang paling berpengaruh dipilih berdasarkan performa model prediksi.

#### 1.4. Batasan Masalah

Penelitian ini memiliki beberapa batasan untuk memastikan fokus penelitian tetap terarah dan sesuai dengan tujuan yang telah ditetapkan. Adapun batasan-batasan dalam penelitian ini adalah sebagai berikut:

1. Fokus penelitian ini berpusat pada penerapan algoritma *Linear Discriminant Analysis* sebagai metode prediksinya dan sekaligus membandingkan metode LDA dengan XGBoost sebagai metode pembandingnya.
2. Penelitian Hanya Mencakup pada 11 sektor industri yang telah diklasifikasikan oleh Bursa Efek Indonesia melalui *Indonesia Stock Exchange Industrial Classification (IDX-IC)*.
3. Penelitian ini akan menggunakan data laporan keuangan perusahaan sebagai sumber informasi untuk melakukan prediksi sektor industri.
4. Data yang digunakan adalah semua daftar perusahaan dari 11 sektor industri Bursa Efek Indonesia melalui *Indonesia Stock Exchange Industrial Classification (IDX-IC)* dari rentang waktu 2010-2022.

#### 1.5. Metode Penelitian

Metode yang dilakukan dalam pengerjaan tugas akhir ini adalah sebagai berikut:

1. Studi Literatur

Melakukan kajian dan mempelajari jurnal serta literatur terkini yang berkaitan dengan metode LDA dan XGBoost. Literatur yang telah dipelajari akan di seleksi berdasarkan relevansinya terhadap pengaplikasian model . Pemilihan Literatur dan jurnal yang relevan akan menjadi pemahaman dasar pada penelitian ini dalam membangun model LDA dan XGBoost untuk memprediksi sektor - sektor industri saham berdasarkan data laporan keuangan perusahaan.

## 2. Pengumpulan Dataset

Dataset laporan keuangan perusahaan yang diambil secara mandiri dan bertahap yang telah terdaftar di Bursa Efek Indonesia (BEI) dalam rentang waktu 2010-2022. Data dikumpulkan dari *Stockbit* dengan memerhatikan fitur-fitur yang dipakai yaitu *Cash, Receivables, Inventory, Total Current Assets, Net Plant & Intangibles, Accounts Payable, Total Debt in Current Liabilities, Unearned Revenue*. Selanjutnya dataset akan dikelompokkan berdasarkan kode sektor industrinya dan tahun laporan keuangannya.

## 3. Pembangunan Model

Dataset yang telah dikumpulkan akan diolah dengan melakukan pra-pemrosesan data, scaling data dan splitting data dengan membagi data menjadi data testing dan data training. data training akan dilakukan *oversampling* untuk mengatasi ketidakseimbangan dataset yang akan dilatih. Selanjutnya membangun Model LDA dan XGBoost serta mengevaluasi dari kinerja kedua melalui *classification report* dan *confusion matrix*.

## 4. Implementasi dan Hasil Pengujian

Implementasi dilakukan dengan menguji model yang telah dibangun, dan dievaluasi dengan melihat hasil *classification report* dan *confusion matrix*.. Hasil pengujian akan dianalisis berdasarkan perbandingan efisiensi model yang menggunakan oversampling dan model yang tidak menggunakan oversampling serta menganalisis perbandingan dari kinerja kedua model untuk melihat perbandingan efisiensi kinerja model .

## 5. Kesimpulan

Penelitian ini ditutup dengan penarikan kesimpulan berdasarkan hasil analisis yang dilakukan pada tahap selanjutnya. Kesimpulan mencakup hasil akurasi yang dihasilkan model LDA dan XGBoost terhadap fitur yang dipilih terhadap akurasi prediksi. Selain itu, penelitian ini memberikan rekomendasi untuk penelitian lanjutan. Tahap ini diakhiri dengan penulisan laporan tugas akhir.

### 1.6. Jadwal Pelaksanaan

Berikut adalah jadwal pelaksanaan Tugas Akhir :

Tabel 1.2. Contoh Jadwal Pelaksanaan Tugas Akhir.

No	Deskripsi Tahapan	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6
1	Studi Literatur						
2	Pengumpulan Data						
3	Eksplorasi Data						
4	Impelementasi Sistem						
5	Analisis Hasil Pengujian						
6	Penyusunan Laporan/Buku TA						