BAB II TINJAUAN PUSTAKA

2.1. Landasan Teori

2.1.1. WordPress dan Plugin

WordPress adalah media *Content Management System* (*CMS*) berbasis web bersifat *open-source* yang berbasis bahasa PHP dan MySQL. Keunggulan utama WordPress terletak pada ketersediaan plugin dalam jumlah besar yang memungkinkan kustomisasi serta penambahan fitur secara lebih bervariasi (Farmana & Yasin, 2022).

Plugin WordPress adalah modul *software* tambahan yang berinteraksi dengan inti WordPress melalui mekanisme *hooks* dan *filters* untuk menyediakan kustomisasi dan penambahan fitur. Namun, tingginya angka penggunaan plugin serta sifatnya yang *open-source* menyebabkan kualitas dan keamanan tidak semuanya terjamin. Sering kali, aspek popularitas lebih diutamakan daripada keamanan, sehingga plugin menjadi salah satu kontributor utama terhadap kerentanan keamanan pada platform WordPress (Ruohonen, 2019).

2.1.2. Database Kerentanan Wordfence

Database kerentanan Wordfence merupakan kumpulan data terstruktur yang berisi informasi terkait celah keamanan (*vulnerability*) pada ekosistem WordPress, khususnya plugin, tema, dan inti sistem. Database ini disediakan oleh Wordfence sebagai bagian dari layanan intelijen keamanannya (*threat intelligence*). Di dalamnya terdapat berbagai detail penting seperti nama plugin yang terdampak, jenis kerentanan (misalnya *SQL Injection*, *Cross-Site Scripting*, atau *Remote Code*

Execution), tingkat keparahan, versi yang terpengaruh, serta status perbaikan (*patched* atau belum). Informasi tersebut diperoleh melalui proses analisis keamanan, pelaporan dari peneliti keamanan, serta pemantauan aktif terhadap ancaman yang beredar. Database ini banyak dimanfaatkan dalam penelitian maupun implementasi sistem keamanan karena menyediakan sumber data yang kredibel dan terus diperbarui, sehingga dapat digunakan sebagai acuan dalam mengidentifikasi, menganalisis, dan memprediksi potensi kerentanan pada plugin WordPress secara sistematis (Wordfence Intelligence, 2026).

2.1.3. WordPress Plugin Metadata

Metadata plugin merupakan informasi non-kode yang tersedia di repositori WordPress. Metadata sering digunakan sebagai indikator kualitas dan keamanan plugin karena mudah diperoleh dan tidak memerlukan analisis kode statis. Plugin populer dengan basis instalasi besar cenderung memiliki lebih banyak kerentanan karena terekspos lebih lama (Ruohonen, 2019). Jumlah unduhan dan *rating* menunjukkan adanya korelasi antara metadata dan kerentanan meskipun hasilnya menunjukkan korelasi lemah (Kasturi dkk., 2022).

2.1.4. *Data Mining*

Data mining adalah proses untuk mengidentifikasi pola, hubungan, tren, atau model yang bermakna dari kumpulan data berukuran besar melalui penerapan teknik statistik, kecerdasan buatan, pembelajaran mesin (*machine learning*), serta sistem basis data. Proses ini bertujuan mengolah data mentah menjadi informasi

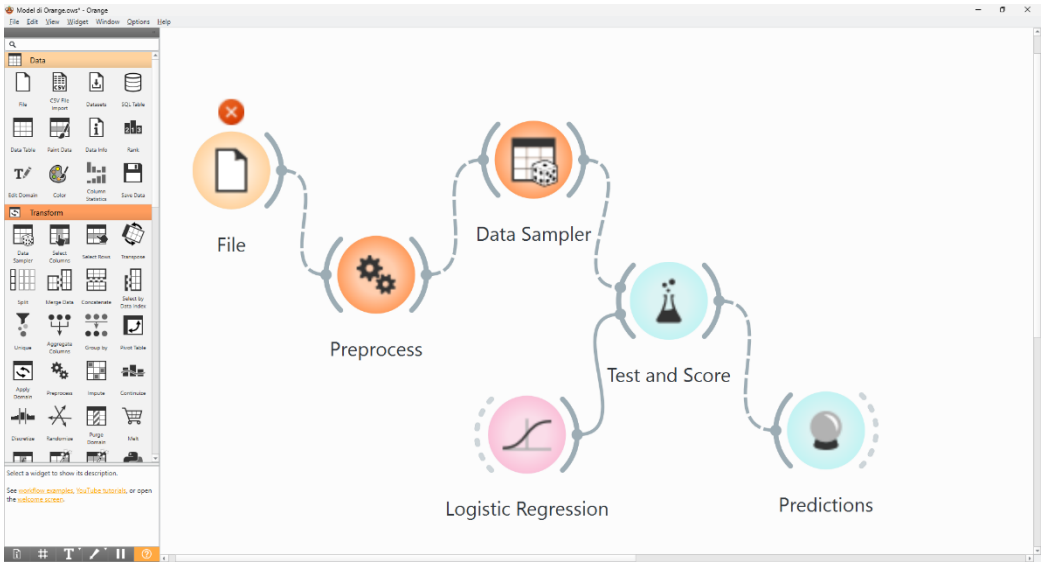
atau pengetahuan yang bernilai guna mendukung pengambilan keputusan dan pemodelan prediktif (Rivaldo dkk., 2024).

Kerangka umum *data mining* menurut (Schröer dkk., 2021) adalah sebagai berikut:

1. *Business/Data Understanding*
2. *Data Preparation*
3. *Modelling*
4. *Evaluation*
5. *Deployment*

2.1.5. Orange

Orange digunakan dalam berbagai penelitian untuk klasifikasi maupun deteksi dalam bentuk prediksi. Contoh penggunaan Orange adalah untuk membandingkan performa algoritma klasifikasi dalam mengklasifikasikan jenis kendaraan, hal ini menunjukkan kemampuan Orange menangani penelitian untuk klasifikasi multi-kelas menggunakan berbagai algoritma *machine learning* (Pranadjaya dkk., 2024). Contoh lainnya, Orange juga telah diterapkan untuk prediksi hasil medis yaitu prediksi penderita diabetes. Penelitian tersebut memperlihatkan bahwa Orange mampu menghasilkan model prediktif yang akurat dengan proses visual dan interaktif (Hartanto & T, 2023).



Gambar 2.1 Halaman Awal *Tool* Orange dengan Contoh

Seperti pada gambar 2.1, keunggulan Orange dapat dilihat, yaitu alat ini memiliki prinsip interaksi kognitif sehingga mudah dipahami. Selain itu, keunggulan Orange lain-nya adalah tingkat efektivitas yang tinggi dalam menyampaikan konsep analisis data melalui visualisasi dan *workflow* interaktif (Dobesova, 2024).

2.1.6. Algoritma *Machine Learning* yang Digunakan

2.1.5.1.Logistic Regression

Logistic Regression merupakan algoritma klasifikasi probabilistik terutama dalam permasalahan klasifikasi biner, di mana variabel target hanya memiliki dua kemungkinan nilai yang dinyatakan sebagai 0 dan 1. Nilai 0 merepresentasikan kondisi tidak terjadinya suatu peristiwa atau kegagalan, sedangkan nilai 1 menunjukkan terjadinya peristiwa atau keberhasilan. Logistic Regression digunakan untuk memodelkan peluang variabel target bernilai 1, sehingga

mendukung proses prediksi dan analisis terhadap hasil biner berdasarkan data yang tersedia (Dey dkk., 2025).

Rumus model Logistic Regression menurut (Dey dkk., 2025) disesuaikan dengan kebutuhan penelitian ini.

$$P(y = 1 \mid x) = \frac{1}{1 + e - (\beta 0 + \beta 1x1 + \beta 2x2 + \cdots + \beta nxn)} \quad (2.1)$$

Di mana:

1. $P(y = 1 \mid x)$ adalah probabilitas plugin rentan,
2. β_0 adalah konstanta,
3. β_i adalah koefisien fitur, dan
4. x_i adalah fitur metadata.

2.1.5.2.Naïve Bayes

Naïve Bayes merupakan salah satu algoritma klasifikasi dalam *machine learning* yang termasuk ke dalam pendekatan *supervised learning*. Algoritma ini didasarkan pada Teorema Bayes yang diperkenalkan oleh Thomas Bayes untuk menghitung probabilitas suatu kejadian. Melalui pendekatan probabilistik, Naïve Bayes mampu merepresentasikan ketidakpastian model dengan menggunakan nilai probabilitas. Dalam konteks klasifikasi, tujuan utama metode ini adalah menghasilkan pemetaan terbaik antara data baru dan kelas-kelas yang telah ditentukan dalam suatu domain permasalahan tertentu (Islamanda & Sibaroni, 2024).

Persamaan matematis yang digunakan dalam metode Naïve Bayes menurut (Islamanda & Sibaroni, 2024) disajikan berdasarkan kebutuhan penelitian ini

$$P(C \mid X) = \frac{P(X \mid C)P(C)}{P(X)} \tag{2.2}$$

Di mana:

1. C adalah kelas (rentan atau tidak rentan), dan
2. X adalah vektor fitur metadata.

2.1.5.3.Random Forest

Random Forest adalah algoritma *ensemble learning* yang membangun sejumlah besar pohon keputusan (*decision trees*) secara independen dan mengombinasikan hasil prediksinya untuk menghasilkan keputusan akhir. Setiap pohon dalam Random Forest dilatih menggunakan sampel data yang diambil secara acak dengan metode *bootstrap* serta subset fitur yang dipilih secara acak pada setiap proses pemisahan node (Breiman, 2001).

Untuk kebutuhan klasifikasi, prediksi Random Forest ditentukan melalui mekanisme *majority voting*, yaitu kelas yang paling sering diprediksi oleh seluruh pohon keputusan. Secara matematis, prediksi akhir dirumuskan sebagai:

$$\acute{y} = mode\{h1(x), h2(x), \dots, hk(x)\} \tag{2.3}$$

di mana $h_k(x)$ merupakan hasil prediksi dari pohon keputusan ke-k.

2.1.7. SHAP (SHapley Additive exPlanations)

SHAP merupakan metode interpretasi model *machine learning* yang diperkenalkan oleh Lundberg dan Lee (2017) melalui pendekatan berbasis teori permainan. Metode ini bertujuan untuk menjelaskan prediksi model dengan memberikan nilai kontribusi pada setiap fitur terhadap hasil prediksi. SHAP

menghitung kontribusi fitur menggunakan konsep Shapley Value, di mana setiap fitur dianggap sebagai *player* yang berkontribusi dalam menghasilkan output model. Nilai SHAP menunjukkan seberapa besar pengaruh suatu fitur dalam meningkatkan atau menurunkan nilai prediksi dibandingkan dengan nilai dasar (*baseline*). Pendekatan ini memiliki keunggulan dalam memberikan interpretasi yang konsisten dan akurat, serta mampu menggabungkan berbagai metode interpretasi model ke dalam satu kerangka kerja yang terpadu (Lundberg & Lee, 2017).

2.2. Penelitian Terkait (*State of the Art*)

Penelitian terkait mendefiniskan posisi dari penelitian ini dibandingkan dengan penelitian sebelumnya sehingga memberikan peran yang jelas bagi penelitian yang akan dilaksanakan dijelaskan pada Tabel 2.1.

Tabel 2.1 Penelitian Terkait (State of the Art)

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
1.	(Khan dkk., 2025)	*Cyber Threat and Vulnerability Classification Using NLP and Machine Learning Techniques on Text-Based Security Data*	Mengklasifikasikan ancaman dan kerentanan keamanan siber secara otomatis dengan memanfaatkan teknik *Natural Language Processing* (*NLP*) dan	Pemanfaatan pendekatan *machine learning* untuk mengklasifikasikan aspek keamanan siber, khususnya yang berkaitan dengan ancaman dan kerentanan.	**Penelitian terkait:** Menggunakan tipe data berbasis teks dengan teknik *Natural Language Processing* (*NLP*) dikombinasi dengan *machine learning*.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
			machine learning terhadap data keamanan berbasis teks.		**Penelitian ini:** Menggunakan jenis data berupa metadata dan menekankan penelitian pada perbandingan algoritma yang dipakai dalam mendeteksi kerentanan.
2.	(Ajayi dkk., 2025)	*Leveraging VAE-Derived Latent Spaces for Enhanced Malware Detection*	Meningkatkan akurasi deteksi *malware* dengan memanfaatkan fitur laten hasil *Variational*	Perbandingan performa alogritma Naïve Bayes, Logistic Regression, dan Random Forest dalam konteks keamanan.	**Penelitian terkait:** Memiliki tujuan utama untuk meningkatkan kinerja deteksi *malware* kemudian

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
		with Machine Learning Classifiers	*Autoencoder (VAE)* sehingga proses klasifikasi *machine learning* menjadi lebih efektif dan efisien.		membandingkan algoritma tersebut. Data yang dipakai adalah data perilaku *malware*. **Penelitian ini:** Memiliki tujuan untuk membandingkan tiga model *machine learning* pada objek prediksi plugin rentan WordPress. Data yang dipakai adalah metadata plugin WordPress.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
3.	(Pranadjaya dkk., 2024)	Perbandingan Algoritma *Machine Learning* menggunakan Orange Data Mining untuk Klasifikasi Jenis Kendaraan pada Sistem Tilang Digital	Melakukan perbandingan dan analisis terhadap algoritma Logistic Regression, Support Vector Machine (SVM), dan Neural Network (NN) dalam menyelesaikan permasalahan klasifikasi kendaraan pada sistem tilang digital.	Penggunaan algoritma Logistic Regression sebagai salah satu model yang dipakai pada penelitian.	**Penelitian terkait:** Model *machine learning* dipakai untuk proses klasifikasi menggunakan jenis data berupa citra. **Penelitian ini:** Model *machine learning* dipakai untuk proses prediksi menggunakan jenis data yaitu metadata.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
4.	(Watanabe dkk., 2023)	*Machine learning Based Prediction of Vulnerability Information Subject to a Security Alert*	Memprediksi dan mengklasifikasikan informasi kerentanan perangkat lunak yang relevan terhadap suatu *security alert* dengan memanfaatkan teknik machine learning.	Pendekatan *machine learning* Logistic Regression dan Random Forest berperan untuk memprediksi dan mengklasifikasikan kerentanan perangkat lunak yang dipakai sebagai pendukung pengambilan keputusan keamanan dan mitigasi risiko berbasis data.	**Penelitian terkait:** Berfokus pada prediksi informasi keamanan dengan memanfaatkan data *security alert*. **Penelitian ini:** Berfokus pada prediksi plugin rentan pada WordPress menggunakan tipe data metadata.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
5.	(Putri dkk., 2023)	Komparasi Algoritma K-NN, Naïve Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir	Melakukan perbandingan terhadap algoritma K-Nearest Neighbor (K-NN), Naïve Bayes Classifier (NBC), dan Support Vector Machine (SVM) dalam menentukan metode yang paling efektif untuk memprediksi tingkat	Membandingkan dan menggunakan algoritma *machine learning* Naïve Bayes untuk prediksi.	**Penelitian terkait:** Menggunakan data akademik dan administratif untuk proses prediksi. **Penelitian ini:** Menggunakan metadata untuk melakukan prediksi.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
			kelulusan mahasiswa pascasarjana.		
6.	(Hartanto & T, 2023)	Implementasi Orange Data Mining untuk Prediksi Penderita Diabetes	Melakukan uji analisis performa *tool* Orange untuk prediksi potensi penderita diabetes.	Penggunaan *data mining* dengan tujuan mengevaluasi kemampuan *tool* tersebut.	**Penelitian terkait:** Uji penelitian pada bidang kesehatan yang memerlukan kebutuhan data medis pasien untuk deteksi potensi terkena diabetes. **Penelitian ini:** Berfokus pada plugin WordPress dengan kebutuhan data seperti metadata untuk

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
					deteksi plugin rentan secara *machine learning*.
7.	(Pham dkk., 2023)	*A Metadata-Based Approach for Research Discipline Prediction Using Machine Learning Techniques and Distance Metrics*	Mengembangkan metode prediksi disiplin penelitian secara otomatis dengan memanfaatkan metadata penelitian menggunakan teknik *machine learning* dan metrik jarak	Pemanfaatan metadata sebagai fitur utama dalam pendekatan *machine learning* untuk melakukan proses prediksi atau klasifikasi	**Penelitian terkait:** Berfokus pada prediksi disiplin keilmuan menggunakan metadata akademik. **Penelitian ini:** Berfokus untuk deteksi plugin rentan pada WordPress menggunakan data berupa metadata perangkat lunak.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
8.	(Kasturi dkk., 2022)	*Mistrust Plugins You Must: A Large-Scale Study Of Malicious Plugins In WordPress Marketplaces*	Mengumpulkan Data plugin Berbahaya di WordPress *Marketplace*	Membedakan plugin yang aman dan berbahaya pada WordPress	**Penelitian terkait:** Melakukan studi plugin berbahaya skala besar tanpa menggunakan machine learning. **Penelitian ini:** Mengevaluasi kemampuan *machine learning* untuk melakukan deteksi plugin rentan WordPress
9.	(Adetunji dkk., 2021)	*House Price Prediction using Random Forest*	Mengeksplorasi penggunaan teknik *machine learning*	Model algoritma *machine learning* untuk prediksi	**Penelitian terkait:** Memanfaatkan model Random Forest dan

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
		Machine Learning Technique	Random Forest untuk prediksi harga rumah.	yang dipakai adalah Random Forest	melakukan optimasi akurasi untuk prediksi dengan output berupa nilai kontinu. Data yang dipakai adalah data harga rumah. **Penelitian ini:** Membandingkan performa model Random Forest pada proses prediksi dengan nilai output berupa label. Data yang dipakai adalah metadata.

NO	Peneliti	Judul	Tujuan	Persamaan Penelitian Terkait dengan Penelitian yang Dilakukan	Perbedaan Penelitian Terkait dengan Penelitian yang Dilakukan
10.	(Murphy dkk., 2021)	*Plugins to Detect Vulnerable Plugins: An Empirical Assessment of the Security Scanner Plugins for WordPress*	Menguji kemampuan 11 plugin WordPress gratis yang berfungsi sebagai *security scanner* deteksi plugin-plugin WordPress berbahaya.	Melakukan percobaan kemampuan berapa metode deteksi plugin rentan dan membandingkannya.	**Penelitian terkait:** Membandingkan performa plugin yang tersedia pada WordPress untuk deteksi kerentanan. **Penelitian ini:** Membandingkan model *machine learning* dalam deteksi plugin WordPress rentan.

2.3. Matriks Penelitian

Matriks penelitian berfungsi untuk membandingkan penelitian-penelitian terdahulu dengan penelitian yang akan dilakukan. Penyusunan matriks ini berdasarkan pada berbagai sumber jurnal yang telah dianalisis dalam *State of the Art*. Tabel 2.2 menyajikan perbandingan yang menunjukkan persamaan dan perbedaan antara penelitian yang diusulkan dengan penelitian-penelitian terkait.

Tabel 2.2 Matriks Penelitian

No	Penulis dan Tahun	Ruang Lingkup Penelitian						
		Alat	Jenis Data	Model			Objek Klasifikasi	
		Orange	*Metadata*	*Logistic Regression*	*Naïve Bayes*	*Random Forest*	*Plugin Rentan WordPress*	*Lainnya*
1.	(Khan dkk., 2025)			✓	✓	✓		✓
2.	(Ajayi dkk., 2025)			✓	✓	✓		✓
3.	(Pranadjaya dkk., 2024)	✓		✓				✓

No	Penulis dan Tahun	Ruang Lingkup Penelitian						
		Alat	Jenis Data	Model			Objek Klasifikasi	
		Orange	*Metadata*	*Logistic Regression*	*Naïve Bayes*	*Random Forest*	*Plugin Rentan WordPress*	*Lainnya*
4.	(Watanabe dkk., 2023)			✓		✓		✓
5.	(Putri dkk., 2023)				✓			✓
6.	(Hartanto & T, 2023)	✓			✓			✓
7.	(Pham dkk., 2023)		✓			✓		✓
8.	(Kasturi dkk., 2022)						✓	
9.	(Adetunji dkk., 2021)					✓		✓
10.	(Murphy dkk., 2021)						✓	
11.	Penelitian ini	✓	✓	✓	✓	✓	✓	

Kebaruan penelitian ini terletak pada implementasi dan analisis kinerja tiga algoritma machine learning, yaitu Logistic Regression, Naïve Bayes, dan Random Forest, untuk memprediksi kerentanan plugin WordPress dengan memanfaatkan metadata plugin sebagai sumber data utama. Berbeda dari penelitian sebelumnya yang umumnya berfokus pada analisis kode statis, deteksi *malware* dengan *machine learning* (Ajayi dkk., 2025), atau penggunaan metode manual (Kasturi dkk., 2022), penelitian ini menekankan penggunaan data non-kode yang mudah diperoleh melalui API resmi WordPress sehingga lebih efisien dan praktis untuk diterapkan. Selain itu, penelitian ini menyajikan evaluasi komprehensif terhadap performa masing-masing algoritma menggunakan metrik evaluasi, sehingga memberikan kontribusi berupa rekomendasi algoritma yang paling efektif untuk prediksi kerentanan plugin WordPress berbasis metadata.