

 2.2  
Impact 7893  
Google Citations Sinta 3  
Current Accreditation[!\[\]\(17413706fd4997a1a4bdf85c6864eee1\_img.jpg\) Google Scholar](#) [!\[\]\(f419710cbe076aa30a9c6c031b5cbe84\_img.jpg\) Garuda](#) [!\[\]\(2726020a4107bdc9042b257034f90eb3\_img.jpg\) Website](#) [!\[\]\(9459655bf14a84f4d775e8d814cca8c9\_img.jpg\) Editor URL](#)

## History Accreditation

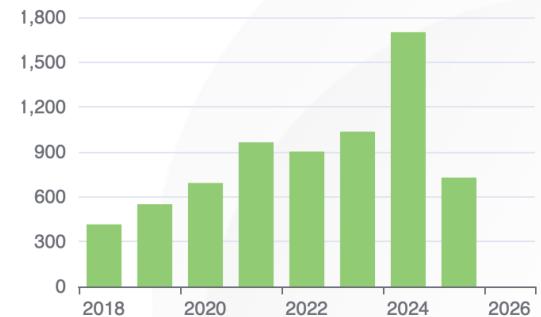
2019 2020 2021 2022 2023 2024 2025 2026 2027



Garuda

Google Scholar

## Citation Per Year By Google Scholar



## Journal By Google Scholar

	All	Since 2021
Citation	7893	6035
h-index	37	29
i10-index	212	159



## Implementasi Majority Voting pada Framework *Cross-Industry Standard Process for Data Mining* untuk Prediksi Kepatuhan Wajib Pajak

Antonius Jonet Binarto<sup>1\*</sup>, Arief Wibowo<sup>2</sup>

<sup>1,2</sup>Universitas Budi Luhur, Indonesia

\*email: 2311601260@student.budiluhur.ac.id

Info Artikel	ABSTRAK
Dikirim: 12 Desember 2024	
Diterima: 16 April 2025	
Diterbitkan: 24 Mei 2025	
<b>Kata kunci:</b>	
<i>Algoritma;</i> <i>Ansambel learning;</i> <i>Machine learning;</i> <i>majority voting;</i> <i>crisp-dm;</i> <i>kepatuhan;</i> <i>rapidminer.</i>	Penelitian ini bertujuan untuk memprediksi kepatuhan Wajib Pajak (patuh atau tidak patuh) menggunakan data perpajakan sebanyak 2.167 baris data. Framework CRISP-DM ( <i>Cross-Industry Standard Process for Data Mining</i> ) digunakan sebagai panduan proses, karena mempunyai kerangka kerja yang terstruktur. Lima algoritma pembelajaran mesin dibandingkan, yaitu Naive Bayes, Support Vector Machine (SVM), Decision Tree, Logistic Regression, dan Deep Learning, dilatih dan diuji menggunakan tools RapidMiner. Untuk meningkatkan akurasi prediksi, metode ansambel <i>majority voting</i> yang merupakan ansambel paling sederhana dan efisien digunakan dengan menggabungkan hasil prediksi dari algoritma-algoritma tersebut dan dievaluasi dan diimplementasikan pada Google Colab menggunakan Python untuk memvalidasi performa pada data baru dan berhasil memberikan akurasi yang lebih stabil dibandingkan model individu. Penelitian ini memberikan kontribusi dalam pengelolaan data perpajakan, khususnya pengambil kebijakan dapat mengoptimalkan penggunaan teknologi untuk meningkatkan efisiensi proses pengawasan dan evaluasi kepatuhan Wajib Pajak. Penelitian ini juga menggarisbawahi pentingnya eksplorasi berbagai algoritma pembelajaran mesin dan ansambel serta parameter lain untuk menghasilkan solusi yang efektif di bidang perpajakan.

### 1 PENDAHULUAN

Pajak merupakan kontribusi wajib yang dibebankan negara kepada warga negaranya tanpa timbal balik langsung, baik perseorangan maupun badan usaha, sebagai sumber pendapatan negara untuk membiayai pengeluaran-pengeluaran negara berdasarkan perundang-undangan dan peraturan yang berlaku [1]. Di Indonesia, pajak merupakan penerimaan utama dan bersifat memaksa [2] dimana hasilnya tidak dapat dinikmati secara langsung, dan merupakan komponen penting dalam pembangunan negara karena merupakan sumber penerimaan negara [3] yang digunakan dalam melakukan pembiayaan dan pembangunan untuk masyarakat banyak. Penerimaan pajak tidak akan dapat tercapai tanpa ada kepatuhan Wajib Pajak dalam melakukan kewajiban mereka, kepatuhan pajak yang tinggi sangat diharapkan dalam hal ini [4]. Kepatuhan pajak dilandasi oleh keyakinan wajib pajak dalam memahami serta menerapkan peraturan perpajakan, hal ini yang dapat memicu perbedaan sikap dari wajib pajak dalam melaksanakan kewajiban perpajakan mereka menurut peraturan yang berlaku [5]. Kurangnya kepatuhan pajak di sebagian besar masyarakat Indonesia merupakan hambatan bagi kemajuan dan pembangunan negara [6] karena akan menghambat juga penerimaan negara sebagai modal pembiayaan dan pembangunan. Untuk itu, perlu suatu tindakan untuk dapat meningkatkan kesadaran Wajib Pajak yang merupakan faktor yang datang dari masing-masing wajib pajak untuk memenuhi kewajiban perpajakannya dengan sukarela dan tanpa paksaan [7]. Dengan memanfaatkan kekuatan komputasi dan algoritma, pemerintah dapat melakukan analisis mendalam terhadap data perpajakan yang sangat besar dan kompleks, sehingga mampu mengungkap pola-pola yang tersembunyi dan sulit dideteksi

atas penghindaran pajak. Hal ini memungkinkan pemerintah untuk mengambil tindakan yang lebih tepat sasaran dalam rangka meningkatkan kepatuhan pajak [8].

Studi empiris pada beberapa Kantor Pelayanan Pajak menunjukkan bahwa penggunaan teknologi dalam bidang perpajakan akan memudahkan Wajib Pajak dalam melaksanakan kewajiban mereka sehingga dapat meningkatkan kepatuhan dalam membayar pajak [9]. Teknologi akan memudahkan dalam pengolahan data yang akan dijadikan landasan dalam menentukan suatu keputusan [10], sementara pada pandangan lain menyebutkan bahwa penggunaan teknologi dalam pengelolaan perpajakan dapat mengurangi potensi penyimpangan perpajakan yang mungkin dapat dilakukan oleh Wajib Pajak [11]..

Penelitian sejenis tentang permasalahan kepatuhan wajib pajak [8] memberikan hasil bahwa penelitian ini menunjukkan bahwa penggunaan algoritma dan kecerdasan buatan pada pemerintahan sangat berpengaruh dalam meningkatkan kepatuhan wajib pajak. Hasil penelitian menunjukkan bahwa 75,30% tingkat kepatuhan wajib pajak di KPP Pratama Jakarta Mampang Prapatan dapat dijelaskan oleh penggunaan teknologi ini. Sisanya, 24,70%, dipengaruhi oleh faktor lain yang tidak diteliti dalam penelitian ini..

Penelitian lain yang menggunakan metodology CRISP-DM [12] menceritakan tentang tujuan yang terkait dengan pengembangan CRISP-DM adalah untuk menciptakan kesan bahwa penambangan data adalah praktik rekayasa yang mapan dengan hasil bahwa CRISP-DM sangat berorientasi pada solusi desain dibandingkan masalah desain.

Penelitian lainnya tentang penggunaan *majority voting* [13] membuktikan bahwa penggunaan ansambel *majority voting* yang digabungkan dengan model CNN dihasilkan sensitivitas 99,4, spesifitas 96,7, AUC sebesar 98,3, dan akurasi 98,5 dalam mengklasifikasikan sel darah kanker dari sel normal, metode yang diusulkan dapat dicapai akurasi tinggi tanpa campur tangan operator dalam penentuan fitur sel. Penelitian terkait tentang kepatuhan Wajib Pajak yang lain yaitu penulis membuat kajian tentang algoritma jaringan syaraf tiruan untuk mendeteksi secara dini kepatuhan wajib pajak orang pribadi [14] yang membandingkan antara model algoritma Multiplayer Perceptron (MLP) dan Support Vector Machine (SVM) dimana didapatkan bahwa MLP mendapatkan nilai lebih tinggi yaitu tingkat akurasi 82,15 persen dan nilai AUC 0,766.

Pada penelitian ini menggunakan data perpajakan dari tempat penulis bekerja, yaitu data tindakan perpajakan atas semua jenis pajak baik orang pribadi maupun badan dalam lingkup satu kantor pajak sejumlah 2.167 baris data, dengan penerapan metodology CRISP-DM atau *Cross-Industry Standard Process for Data Mining* diharapkan akan mendapatkan melakukan proses yang terstruktur dan dapat meningkatkan pemahaman antar pemangku kepentingan [15] karena metode ini merupakan pendekatan yang matang dan sudah diterima secara luas dalam praktik penambangan data [16], dengan membandingkan secara individu hasil dari prediksi tiap algoritma yang dipilih dalam menentukan tingkat kepatuhan Wajib Pajak, dan diikuti dengan metode ansambel pengambilan suara terbanyak atau *majority voting* untuk hasil akhir agar dapat memprediksi lebih akurat atas model yang dipilih karena metode ini adalah algoritma yang paling sederhana dan efisien tetapi sangat kuat dan terbukti dalam memberikan prediksi [13]. Sebagai perbandingan dengan penelitian sebelumnya yang menggunakan algoritma individual untuk memprediksi kepatuhan wajib pajak, penelitian ini mengadopsi pendekatan ansambel *majority voting* yang mampu meningkatkan stabilitas prediksi dengan menggabungkan kekuatan dari beberapa algoritma machine learning. Hasil evaluasi menunjukkan bahwa metode ini menghasilkan prediksi yang lebih akurat dibandingkan dengan pendekatan tunggal, sehingga dapat menjadi solusi yang lebih andal dalam analisis kepatuhan wajib pajak berbasis data.

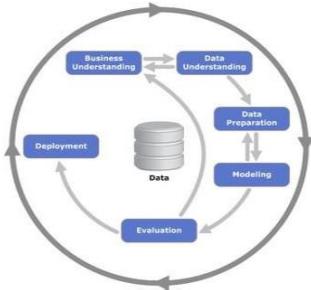
## 2 METODE PENELITIAN

Pada penelitian ini, penulis menggunakan kerangka kerja CRISP-DM (*Cross Industry Standard Process Model For Data Mining*) karena mempunyai kerangka kerja yang jelas dan terstruktur dengan tahapan yang sangat ideal dan komprehensif, sementara untuk pemodelan data digunakan tools RapidMiner yang sudah terbukti mampu untuk mengolah data dalam skala kecil maupun besar dan sangat *user friendly* menurut penulis dengan beberapa algoritma yang digunakan sebagai bentuk pembanding yang nantinya akan divalidasi pada lingkup simulasi menggunakan tools *Google Colab* yang memungkinkan dengan skrip program sederhana dengan bahasa *python* untuk menjalankan perintah prediksi dengan mengadopsi ansambel *majority voting* yang

merupakan ansambel paling efisien dan sederhana dalam penerapannya dengan tujuan akan mendapatkan prediksi yang paling akurat.

## 2.1 CRISP-DM

Framework atau kerangka kerja *Cross Industry Standard Process Model For Data Mining* (CRISP DM) (Gambar 1) merupakan suatu proses kerja dalam memecahkan permasalahan yang terjadi, dengan tahapan mulai dari *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment* [17].



Gambar 1. Alur CRISP-DM

## 2.2 Business Understanding

Tahap ini merupakan tahap pertama yang perlu dilakukan untuk memahami masalah dan tujuan kebutuhan suatu bisnis dalam hal ini meningkatkan kepatuhan Wajib Pajak sehingga dapat ditentukan rencana dan strategi guna mencapai tujuan tersebut.

## 2.3 Data Understanding

Tahapan ini mulai dari pengumpulan data sampai pemahaman karakteristik tentang data tindakan perpajakan dan pemenuhan kewajiban oleh Wajib Pajak yang digunakan sangat diperlukan pada tahap ini, analisis deskriptif dilakukan untuk memahami karakteristik data, menemukan pola awal, dan mengidentifikasi masalah kualitas data seperti nilai yang hilang atau outlier.

## 2.4 Data Preparation

Tahap persiapan data melibatkan pembersihan, transformasi, dan konstruksi data sehingga siap untuk digunakan dalam pemodelan. Kegiatan yang umum dilakukan pada tahap ini meliputi penanganan nilai yang hilang, normalisasi data, dan pembuatan fitur atau attribute baru.

## 2.5 Modeling

Tahap ini merupakan inti dari proses, berbagai algoritma pembelajaran mesin (*machine learning*) diterapkan pada data yang telah disiapkan untuk membangun model atau pola prediksi, dimana pemilihan algoritma yang tepat sangat bergantung pada jenis masalah dan karakteristik data. Model yang didapatkan pada proses modeling ini nantinya akan digunakan dalam proses tahap selanjutnya. Pemilihan algoritma machine learning dalam penelitian ini didasarkan pada karakteristik dataset yang digunakan serta keunggulan masing-masing algoritma dalam menangani pola data yang berbeda. Penggunaan metode ansambel majority voting bertujuan untuk mengoptimalkan prediksi dengan menggabungkan keunggulan dari berbagai pendekatan tersebut.

### 1) Algoritma Naïve Bayes

Naive Bayes adalah algoritma klasifikasi sederhana dengan pendekatan probabilitas [18]. Prinsip utamanya adalah teorema Bayes, dengan asumsi bahwa fitur bersifat independen satu sama lain, (persamaan (1)) menghitung probabilitas posterior  $(C|X)$  dari probabilitas sebelumnya  $P(C)$ ,  $P(X|C)$

yang merupakan probabilitas sebelumnya dari X di mana C adalah hipotesis dan X adalah kumpulan data yang diberikan.

2) Algoritma SVM

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. SVM bekerja dengan mencari *hyperplane* yang memisahkan data menjadi kategori yang berbeda dengan margin maksimum. SVM adalah algoritma umum yang didasarkan pada batasan risiko terjamin dari teori pembelajaran statistik yaitu yang disebut prinsip minimisasi risiko struktural [19]. Prinsip penghitungan *hyperlane* yaitu vektor bobot dikalikan dengan vektor fitur dan dijumlahkan dengan bias.

3) Algoritma Decision Tree

Adalah metode pembelajaran mesin yang digunakan untuk klasifikasi dan regresi dengan membangun struktur berbentuk pohon dimana setiap node internal mewakili tes pada fitur, cabang mewakili hasil tes, dan daun mewakili keputusan atau prediksi. Pada dasarnya algoritma ID3 membagi data menjadi dua kelompok berdasarkan atribut homogenitas (A) yang ada pada data (S) tersebut, dengan cara mengukur suatu bilangan yang disebut entropi [20]. Dari entropi dapat diketahui pengurang pada gain atau gini setelah dibagi dengan dataset berdasarkan fitur untuk mendapatkan hasil seberapa baik fitur tersebut.

4) Algoritma Logistic Regression

Algoritma pembelajaran mesin yang digunakan untuk memprediksi hasil berbentuk kategori atau klasifikasi dengan memodelkan hubungan antara variabel independen (fitur) dan variabel dependen (kelas) dengan menggunakan fungsi logistik atau sigmoid. dengan kata lain Variabel dependen biasanya hanya terdiri atas dua nilai yang mewakili kemunculan atau tidak adanya suatu kejadian yang biasanya diberi angka 0 atau 1 [21].

5) Algoritma Deep Learning

Algoritma ini adalah evolusi dari jaringan saraf tiruan yang mampu melakukan pembelajaran secara mandiri menyesuaikan parameter-parameternya dan memilih arsitektur jaringan yang paling optimal untuk tugas yang diberikan, tanpa memerlukan data pelatihan yang berlabel [22]. Deep learning dirancang untuk secara otomatis mengekstraksi fitur dari data mentah melalui proses

6) Majority Voting

Merupakan teknik ansambel learning yang paling efisien dan sederhana dalam penggunaannya dengan menggabungkan prediksi dari beberapa model untuk meningkatkan akurasi prediksi. Dalam konteks prediksi kepatuhan wajib pajak, kita dapat menggunakan *majority voting* untuk menggabungkan prediksi dari beberapa model *machine learning* atau algoritma. Dalam pembelajaran ansambel, pemungutan suara mayoritas digunakan khususnya untuk klasifikasi multinomial, di mana setiap model memberikan suaranya satu kelas, dan kelas dengan suara terbanyak dianggap sebagai prediktor akhir [23].

## 2.6 Evaluation

Model yang telah dibangun kemudian dievaluasi untuk mengukur kinerjanya dan hasil prediksi yang didapatkan. Untuk memenuhi standar evaluasi, berbagai metrik evaluasi digunakan seperti akurasi, presisi, *recall*, dan metrik lainnya yang relevan menggunakan skrip bahasa *python* sederhana pada google colab, hasil dengan nilai tertinggi merupakan hal yang ingin dicapai sehingga relevan untuk dapat diterapkan dalam penelitian ini dan bahkan dalam penerapan di lapangan nantinya. Evaluasi akan dilakukan secara individu tiap algoritma dan kemudian akan digunakan ansambel *majority voting* untuk mengetahui peningkatan akurasi dari hasil prediksi akhir.

## 2.7 Deployment

Tahap terakhir adalah penerapan model yang telah teruji ke dalam sistem produksi atau lingkungan bisnis, pada penelitian ini deployment dilakukan pada lingkungan simulasi yang disiapkan, dimana setelah mendapatkan prediksi dari beberapa model, dilakukan voting untuk menentukan prediksi akhir. Model dengan suara terbanyak akan menjadi prediksi akhir. Penggunaan data baru dan skrip program sederhana pada google colab

untuk melakukan deployment pada tiap algoritma diharapkan akan mendapatkan hasil sebelum akhirnya akan digunakan ansambel majority voting untuk mengetahui hasil akhir paling baik pada tahap deployment.

### 3 HASIL DAN PEMBAHASAN

#### 3.1 Data Preparation

##### 1) Pengumpulan Data

Data (Tabel 1) yang didapat merupakan data dari tempat penulis bekerja, dimana data tersebut merupakan data tindakan perpajakan yang sudah pernah dilakukan sebanyak 2.167 baris data,

Tabel 1. Hasil Pengumpulan Data

No	Saldo	Teguran	Paksa	Sita	Keberatan	Nonkeberatan
1.	-87030542	29	1	0	1	2
2.	6500000	13	0	0	0	0
3.	897695868	39	9	0	5	32
4.	0	36	12	0	0	17
5.	3499999	12	11	0	0	0
6.	1200000	14	13	0	0	0
7.	0	26	13	0	0	21
...	...	...	...	...	...	...
...	...	...	...	...	...	...
2167.	417008	11	0	0	0	0

dikelompokkan menjadi beberapa attribute (Tabel 2) yaitu saldo, teguran, paska, sita, keberatan, non-keberatan dan status.

Tabel 2. Struktur Data

No	Attribute	Tipe	Keterangan
1.	Saldo	real	merupakan data pemenuhan tunggakan pajak dari Wajib Pajak
2.	Teguran	integer	data Surat Teguran yang dikeluarkan dalam rangka penagihan pajak
3.	Paksa	integer	data Surat Paksa yang dikeluarkan dalam rangka penagihan pajak
4.	Sita	integer	data Surat Sita yang dikeluarkan dalam rangka penagihan pajak
5.	Keberatan	integer	jumlah pengajuan permohonan keberatan oleh Wajib Pajak
6.	Non Keberatan	integer	jumlah pengajuan permohonan non keberatan oleh Wajib Pajak
7.	Status	polynomial	label kategori PATUH atau TIDAK

##### 2) Pemilihan Data

Data merupakan data hasil olahan dari berbagai sumber tabel, dan atas data bilangan pecahan dikeluarkan dari dataset dalam proses pemilihan data ini, dan dari 2.167 data kemudian dibagi menjadi 1.667 baris data test dan 500 baris data uji.

##### 3) Pelabelan Data

Penentuan label (Tabel 3) digunakan dengan pengamatan dan validasi atas attribute yang terlibat.

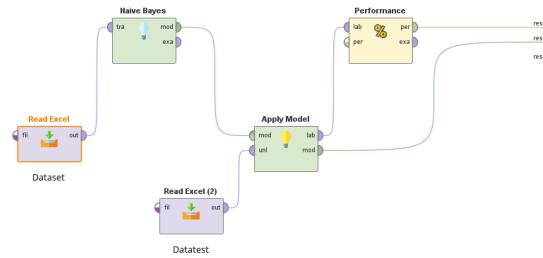
Tabel 3. Dataset

No	Saldo	Teguran	Paksa	Sita	Keberatan	Nonkeberatan	Status
1.	-87030542	29	1	0	1	2	PATUH
2.	6500000	13	0	0	0	0	TIDAK
3.	897695868	39	9	0	5	32	TIDAK
4.	0	36	12	0	0	17	TIDAK
5.	3499999	12	11	0	0	0	TIDAK
6.	1200000	14	13	0	0	0	TIDAK
7.	0	26	13	0	0	21	TIDAK
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
2167.	417008	11	0	0	0	0	TIDAK

#### 3.2 Modelling

### 1) Hasil Modeling Naïve Bayes

Dengan menggunakan tools RapidMiner, karena memudahkan implementasi Naïve Bayes untuk mengolah data tindakan perpajakan dengan atribut numerik atau integer. Setelah penentuan label pada datatest dan data uji, datatest dihubungkan dengan operator Naïve Bayes (Gambar 2) dan dilakukan pemodelan dengan menggabungkan data uji untuk ditampilkan hasilnya.



Gambar 2. Modeling Naïve Bayes

Hasil yang didapatkan dengan modeling ini yaitu akurasi sebesar 51.00%, dengan *class precision* prediksi PATUH sebesar 39.79% dan *class precision* prediksi TIDAK sebesar 87.29% sementara *class recall* yang didapat untuk *true* PATUH adalah 91.02% dan TIDAK sebesar 30.93%.

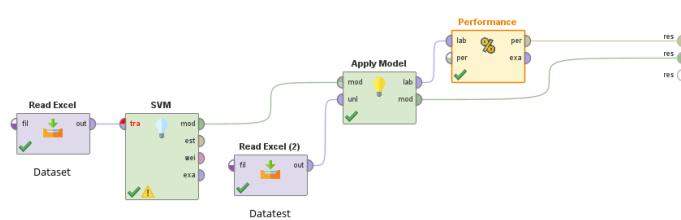
accuracy: 51.00%

	true PATUH	true TIDAK	class precision
pred. PATUH	152	230	39.79%
pred. TIDAK	15	103	87.29%
class recall	91.02%	30.93%	

Gambar 3. Hasil Modeling Naïve Bayes

### 2) Hasil Modeling SVM

Modeling dengan Support Vector Machine dan tools RapidMiner (Gambar 4) kita lakukan untuk mendapatkan akurasi, dengan menggunakan label yang sudah kita identifikasi dengan cara kerja sama dengan pemodelan Naïve Bayes sebelumnya.



Gambar 4. Modeling SVM

Dengan pemodelan menggunakan SVM didapatkan akurasi sebesar 66.60% dengan *class precision* prediksi PATUH sebesar 66.60% dan 0% untuk TIDAK, untuk *class recall true* TIDAK didapatkan 100% dan 0% untuk TIDAK.

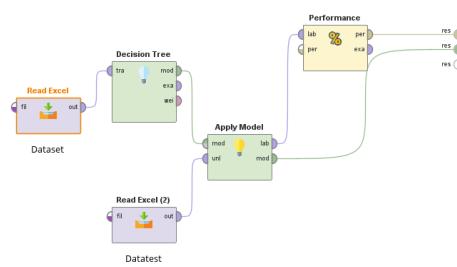
accuracy: 66.60%

	true PATUH	true TIDAK	class precision
pred. PATUH	0	0	0.00%
pred. TIDAK	167	333	66.60%
class recall	0.00%	100.00%	

Gambar 5. Hasil Modeling SVM

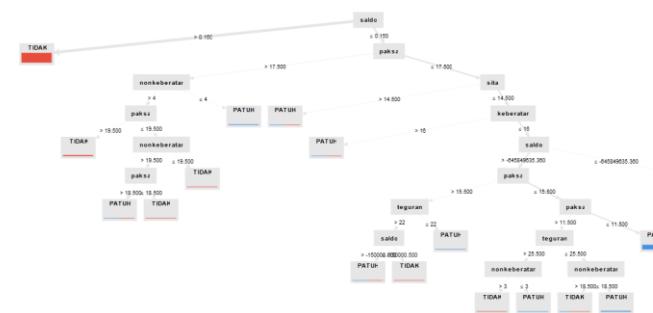
### 3) Hasil Modeling Decision Tree

Pencarian akurasi dengan algoritma Decision Tree pada RapidMiner (Gambar 6), dengan cara yang sama dengan pemodelan sebelumnya, attribute Status menjadi target.



Gambar 6. Modeling Decision Tree

Hasilnya didapatkan suatu pohon keputusan (Gambar 5) dari model di atas.



Gambar 7. Hasil Pohon Keputusan Modeling Decision Tree

Pemodelan ini mendapatkan akurasi sebesar 99.20% dengan *class precision* prediksi PATUH sebesar 98.22% dan TIDAK sebesar 99.70% dan *class recall* yang didapatkan yaitu 99.40% untuk *true* PATUH dan 99.10 untuk *true* TIDAK.

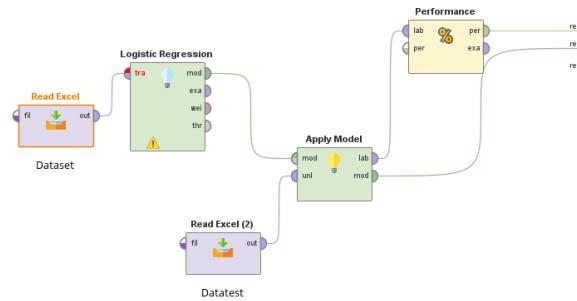
accuracy: 99.20%

	true PATUH	true TIDAK	class precision
pred. PATUH	166	3	98.22%
pred. TIDAK	1	330	99.70%
class recall	99.40%	99.10%	

Gambar 8. Hasil Modeling Decision Tree

#### 4) Hasil Modeling Logistic Regression

Dengan tools yang sama, modeling dilakukan dengan algoritma Logistic Regression (Gambar 9) atas data test dan data uji yang sama.



Gambar 9. Modeling Logistic Regression

Dengan algoritma ini didapatkan akurasi sebesar 68.40% dan *class precision* 61.54% untuk prediksi PATUH serta 68.98% untuk prediksi TIDAK, hasil *class recall* yang didapatkan untuk *true* PATUH sebesar 14.37% dan 95.50% untuk *true* TIDAK.

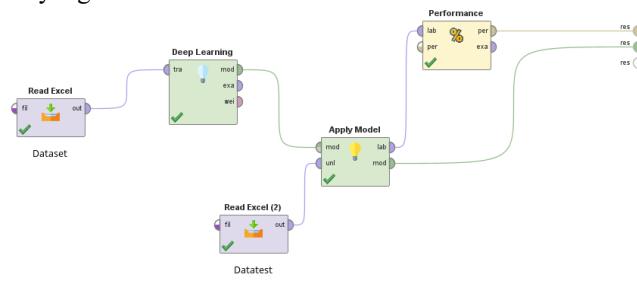
Table View Plot View

accuracy: 68.40%			
	true PATUH	true TIDAK	class precision
pred. PATUH	24	15	61.54%
pred. TIDAK	143	318	68.98%
class recall	14.37%	95.50%	

Gambar 10. Hasil Modeling Logistic Regression

5) Hasil Modeling Deep Learning

Menggunakan operator Deep Learning (Gambar 11), pemodelan dilakukan untuk mendapatkan *performance* atas data yang tersedia.



Gambar 11. Modeling Deep Learning

Hasil yang diperoleh dengan pemodelan ini yaitu akurasi sebesar 75.00% dan untuk *class precision* prediksi PATUH sebesar 63.46 dengan 80.23% untuk prediksi TIDAK, sedangkan *class recall* *true* PATUH didapatkan 59.28% dan *true* TIDAK sebesar 82.88%.

Table View Plot View

accuracy: 75.00%			
	true PATUH	true TIDAK	class precision
pred. PATUH	97	55	63.46%
pred. TIDAK	70	278	80.23%
class recall	59.28%	82.88%	

Gambar 12. Hasil Modeling Deep Learning

### 3.3 Evaluation

Setelah dilakukan pemodelan, tahap evaluasi yang diambil yaitu dengan membandingkan dan mengamati semua hasil modeling tersebut. Evaluasi dilakukan dengan tools google colab dengan data yang sama dan menggunakan bahasa python untuk tiap algoritma yang sudah dimodelkan di atas. Penting untuk dipahami bahwa setiap algoritma memiliki keunggulan dan keterbatasan dalam menangkap pola dari dataset yang digunakan. Oleh karena itu, penelitian ini membandingkan beberapa metrik evaluasi utama, seperti akurasi, precision, recall, dan F1-score, untuk mengukur performa masing-masing model.

Tabel 4. Perbandingan Hasil Evaluasi Model

Uraian	Naïve Bayes	SVM	Decision Tree	Logistic Regression	Deep Learning
Accurasion	34.56%	67.05%	99.08%	67.05%	76.27%
Class Precision	33.17%	67.28%	97.91%	0.00%	70.00%
Class Recall	98.59%	99.31%	99.29%	0.00%	48.00%
Specificity	98.59%	0.70%	99.29%	99.66%	90.00%
F1 Score	49.64%	80.22%	98.60%	0.00%	57.00%

Berdasarkan Tabel 4, algoritma Decision Tree menunjukkan performa terbaik dengan akurasi 99,08%, *class precision* 97,91%, *class recall* 99,29%, *specificity* 99,29%, dan *F1 score* 98,60%. Hasil ini menandakan bahwa Decision Tree sangat efektif dalam menangkap pola pada dataset, meminimalkan kesalahan klasifikasi, dan memberikan prediksi yang andal. Algoritma Deep Learning menempati posisi kedua dengan akurasi 76,27%, *class precision* 70%, *class recall* 48%, *specificity* 90%, dan *F1 score* 57%. Walaupun akurasi cukup tinggi, ketimpangan antara *class recall* dan *specificity* menunjukkan bahwa Deep Learning memiliki kelemahan dalam mendekripsi kategori tertentu pada data. Di sisi lain, algoritma Support Vector Machine (SVM) mencatat performa menengah dengan akurasi 67,05%, *precision* 67,28%, *recall* 99,31%, dan *F1 score* 80,22%, yang menunjukkan keunggulannya dalam *recall* tetapi kelemahan pada *specificity* dengan 0,70%. Naive Bayes menunjukkan hasil yang kurang kompetitif dengan akurasi 34,56%, *precision* 33,17%, dan *F1 score* 49,64%, mengindikasikan asumsi independensi fitur yang tidak cocok dengan data ini. Logistic Regression memiliki hasil terburuk dengan semua metrik selain *specificity* di angka 99,66% dan mencatat nilai 0%. Hal ini kemungkinan disebabkan oleh data yang tidak linier atau preprocessing yang kurang sesuai. Hasil ini menyoroti pentingnya pemilihan algoritma yang tepat berdasarkan karakteristik dataset untuk meningkatkan akurasi dan keandalan prediksi kepatuhan wajib pajak.

Berdasarkan hasil pada Tabel 4, dapat disimpulkan bahwa algoritma Decision Tree menunjukkan performa terbaik dengan akurasi tertinggi, sementara Naïve Bayes memiliki keterbatasan dalam menangani dataset ini karena asumsi independensi fitur yang tidak selalu berlaku. Algoritma Deep Learning memberikan hasil yang menjanjikan, tetapi memiliki tantangan dalam komputasi dan tuning parameter. Oleh karena itu, penggunaan ansambel majority voting dipilih sebagai strategi untuk menggabungkan prediksi terbaik dari berbagai model guna meningkatkan keandalan hasil.

### 3.4 Deployment

Proses deployment dari pemodelan dan pengujian dilakukan masih menggunakan google colab dengan bahasa python untuk mengimplementasikan dan memberikan hasil terhadap data baru yang diberikan pada program. Dengan data baru yang sama dilakukan implementasi untuk setiap model atau algoritma dengan tujuan untuk mendapatkan hasil prediksi dari tiap model.

Tabel 5. Data Baru untuk Deployment

No	Saldo	Teguran	Paksa	Sita	Keberatan	Nonkeberatan
1.	-10000000	11	3	1	0	31

Tabel rangkuman hasil deployment beberapa model seperti di bawah ini:

Tabel 6. Hasil Prediksi Implementasi Model

Model/Algoritma	Hasil Prediksi
Naïve bayes	PATUH
SVM	TIDAK
Decision Tree	PATUH
Logistic Regression	TIDAK
Deep Learning	PATUH

Pada tahap *deployment*, hasil implementasi (Tabel 6) dari lima algoritma prediksi kepatuhan wajib pajak (Naïve Bayes, SVM, Decision Tree, Logistic Regression, dan Deep Learning) menunjukkan variasi hasil yang signifikan. Berdasarkan Tabel 6, algoritma Naïve Bayes, Decision Tree, dan Deep Learning memprediksi status wajib pajak sebagai PATUH. Sebaliknya, algoritma SVM dan Logistic Regression memprediksi status sebagai TIDAK PATUH. Perbedaan ini terjadi karena setiap algoritma memiliki cara berbeda dalam menangkap pola dari data yang diberikan, sehingga menghasilkan prediksi yang tidak selalu konsisten satu sama lain. Oleh karena itu, untuk menentukan hasil akhir yang lebih stabil dan representatif, digunakan metode *majority voting* sebagai pendekatan ansambel.



```

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import VotingClassifier
from sklearn.metrics import accuracy_score
# Buat ensemble voting
ensemble = VotingClassifier(estimators=[('NB', NB), ('SVM', SVM), ('DT', DT), ('LR', LR)], voting='soft')

[11]: ensemble.fit(X_train, y_train)

[12]: y_pred = ensemble.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)

Accuracy: 0.988492926272811

[1]: new_data = {'saldo': 10000000, 'teguran': 11, 'paksa': 3, 'sita': 1, 'keberatan': 0, 'nonkeberatan': 31}
new_data = pd.DataFrame(new_data, index=[0])

# Prediksi status
prediction = ensemble.predict(new_data)
print("Hasil Prediksi voting:", prediction[0])

Hasil Prediksi status: PATUH

```

Gambar 13. Hasil Deploy Ansambel Majority Voting

Ansambel *Majority voting* mempunyai stabilitas yang tinggi ketika digunakan untuk kinerja serupa yang dilakukan oleh beberapa model, dan dapat mengeliminasi atau mengoreksi noise data pada model karena dengan adanya model mayoritas yang lebih kecil noise nya. Sederhana dan simpel dalam penggunaan menjadi pertimbangan penulis dalam mengadopsi ansambel ini, selain karena faktor efisiensi dalam menggabungkan hasil prediksi dari kelima algoritma untuk mencapai keputusan akhir berdasarkan suara terbanyak. Hasilnya, tiga algoritma, yaitu Naïve Bayes, Decision Tree, dan Deep Learning menghasilkan prediksi PATUH, sedangkan dua lainnya, SVM dan Logistic Regression menghasilkan prediksi TIDAK. Dengan demikian, keputusan akhir dari *majority voting* adalah PATUH, karena suara terbanyak diperoleh pada kategori tersebut. Pendekatan ini tidak hanya meningkatkan keandalan hasil prediksi tetapi juga mengurangi pengaruh bias atau kelemahan individu dari satu algoritma tertentu. *Majority voting* memberikan keuntungan signifikan dalam meningkatkan akurasi prediksi secara keseluruhan, terutama ketika setiap algoritma memiliki kekuatan dan kelemahan yang berbeda. Dengan hasil ini, *deployment* sistem prediksi dapat memberikan rekomendasi lebih andal kepada otoritas perpajakan untuk menilai kepatuhan wajib pajak secara lebih efektif.

#### 4 KESIMPULAN

Framework CRISP-DM dalam penelitian ini memungkinkan prediksi kepatuhan wajib pajak secara sistematis, mulai dari pemahaman data hingga evaluasi model. Lima algoritma \*machine learning\* diterapkan, yaitu Naïve Bayes, SVM, Decision Tree, Logistic Regression, dan Deep Learning, dengan hasil prediksi yang bervariasi. Untuk meningkatkan keandalan, metode ansambel \*majority voting\* digunakan, menggabungkan kekuatan tiap algoritma agar hasil akhir lebih akurat. Tantangan ke depan meliputi penanganan ketidakseimbangan data, integrasi data tambahan seperti demografi (usia, pendidikan) dan geografi (lokasi), serta eksplorasi metode ansambel lain seperti \*stacking\* atau \*bagging\*. Selain itu, dataset perlu diperluas agar lebih representatif dan real-time, serta mencakup parameter ekonomi wajib pajak seperti penghasilan, laporan keuangan, dan jumlah harta. Teknologi prediksi ini membantu fiskus dalam mengotomasi analisis perilaku wajib pajak secara real-time, sehingga dapat menyusun rencana kerja serta memperkirakan potensi penerimaan pajak dengan lebih akurat, sekaligus menjadi acuan dalam penelitian dan pemeriksaan perpajakan.

Model prediksi yang dikembangkan dalam penelitian ini dapat memberikan manfaat signifikan bagi otoritas pajak dalam meningkatkan kepatuhan wajib pajak. Dengan mengadopsi sistem berbasis machine learning, otoritas pajak dapat secara otomatis mengidentifikasi kelompok wajib pajak yang memiliki risiko tinggi terhadap ketidakpatuhan, sehingga memungkinkan pengawasan yang lebih efektif dan efisien. Selain itu, model ini dapat digunakan sebagai alat bantu dalam pengambilan keputusan kebijakan perpajakan, seperti penyusunan strategi sosialisasi pajak yang lebih tepat sasaran atau pengembangan kebijakan insentif yang berbasis data. Integrasi model ini dengan sistem informasi perpajakan yang sudah ada juga dapat membantu dalam mempercepat proses verifikasi kepatuhan serta mendukung pengambilan keputusan berbasis data secara real-time.

## REFERENSI

- [1] F. Ristanti, Uswatun Khasanah, and Cris Kuntadi, "Literature Review Pengaruh Penerapan Pajak UMKM, Sosialisasi Perpajakan dan Sanksi Perpajakan terhadap Kepatuhan Wajib Pajak," *J. Ilmu Multidisplin*, vol. 1, no. 2, pp. 380–391, 2022, doi: 10.38035/jim.v1i2.49.
- [2] A. P. Nugraheni, S. N. Sunaningsih, and N. A. Khabibah, "Peran Konsultan Pajak Terhadap Kepatuhan Wajib Pajak," *Jati J. Akunt. Terap. Indones.*, vol. 4, no. 1, p. Editing, 2021, doi: 10.18196/jati.v4i1.9701.
- [3] A. F. Putra, "Kepatuhan Wajib Pajak UMKM: Pengetahuan Pajak, Sanksi Pajak, dan Modernisasi Sistem," *J. Ris. Akunt. Perpajak.*, vol. 7, no. 01, pp. 1–12, 2020, doi: 10.35838/jrap.v7i01.1212.
- [4] Zur'aini, Yahya, and L.M Samsu, "Implementasi Algoritma Naïve Bayes Untuk Memprediksi Tingkat Kepatuhan Wajib Pajak Di Desa Dames Damai," *J. Pengemb. Rekayasa Inform. dan Komput.*, vol. 1, no. 2, pp. 64–77, 2023.
- [5] A. Karina, M. Siska, and B. Saad, "Tingkat Pendapatan, Tarif Pajak, Dan Sanksi Pajak Pengaruhnya Terhadap Kepatuhan Wajib Pajak Orang Pribadi Dengan Pemeriksaan Pajak Sebagai Variabel Moderasi," *Econ. Digit. Bus. Rev.*, vol. 5, no. 1, pp. 356–369, 2024.
- [6] F. Setianingrum, F. Hutabarat, H. Purba, and U. A. Indonesia, "The effect of implementing an e-filing system, the quality of information technology on the level of tax compliance," vol. 7, pp. 979–993, 2024.
- [7] T. K. Hek and P. Gani, "Pengaruh Pengetahuan, Kesadaran dan Kualitas Pelayanan Terhadap Kepatuhan Pajak PBB di Kecamatan Patumbak," *Jesya*, vol. 6, no. 1, pp. 1001–1006, 2023, doi: 10.36778/jesya.v6i1.1002.
- [8] A. Adha, R. Rulinawaty, and F. Madya, "Analisis Pengaruh Pemerintahan Dengan Algoritma Dan Artificial Intelligence (AI) Terhadap Kepatuhan Wajib Pajak Pada Kpp Pratama Jakarta Mampang Prapatan," *J. Econ. Bussines Account.*, vol. 7, no. 4, pp. 10823–10836, 2024, doi: 10.31539/costing.v7i4.11192.
- [9] B. D. Handayani and Y. Friskianty, "Pengaruh Self Assessment System, Keadilan, Teknologi Perpajakan, Dan Ketidakpercayaan Kepada Pihak Fiskus Terhadap Tindakan Tax Evasion," *Account. Anal. J.*, vol. 3, no. 4, pp. 457–465, 2014, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/aaaj>
- [10] A. Ulfa, "Pengaruh Kemungkinan Terdeteksinya Kecurangan, Teknologi dan Informasi Perpajakan, dan Kepercayaan Pada Otoritas Pemerintah Terhadap Penggelapan Pajak (Studi Empiris Pada Orang Pribadi Yang Terdaftar di Kantor Pelayanan Pajak Pratama Senapelan Pekanbaru)," *Jom FEKON*, vol. 2, no. 2, pp. 1–15, 2015, [Online]. Available: <https://jom.unri.ac.id/index.php/JOMFEKON/article/view/9643>
- [11] A. N. Dewi and W. Irawati, "Pengaruh Pemahaman Perpajakan, Keadilan, dan Teknologi Perpajakan terhadap Perilaku Penggelapan Pajak (Studi Empiris pada Wajib Pajak Orang Pribadi Terdaftar di KPP Pratama Serpong)," *Yudishtira J. Indones. J. Financ. Strateg. Insid.*, vol. 2, no. 2, pp. 262–279, 2022.
- [12] U. Kannengiesser and J. S. Gero, "Modelling the Design of Models: an Example Using Crisp-Dm," *Proc. Des. Soc.*, vol. 3, no. July, pp. 2705–2714, 2023, doi: 10.1017/pds.2023.271.
- [13] M. Ghaderzadeh, A. Hosseini, F. Asadi, H. Abolghasemi, D. Bashash, and A. Roshanpoor, "Automated Detection Model in Classification of B-Lymphoblast Cells from Normal B-Lymphoid Precursors in Blood Smear Microscopic Images Based on the Majority Voting Technique," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/4801671.
- [14] T. Rohana and M. Arifuddin, "Kajian Algoritma Jaringan Syaraf Tiruan untuk Mendeteksi Secara Dini Kepatuhan Wajib Pajak Orang Pribadi," *Semin. Nas. Apl. Teknol. Inf.*, pp. 1–6, 2013.
- [15] A. Massahiro *et al.*, "The evolution of CRISP-DM for Data Science : Methods , Processes and Frameworks," 2024, doi: 10.5753/reviews.2024.3757.
- [16] J. A. Solano, D. J. Lancheros Cuesta, S. F. Umaña Ibáñez, and J. R. Coronado-Hernández, "Predictive models assessment based on CRISP-DM methodology for students performance in Colombia - Saber 11 Test," *Procedia Comput. Sci.*, vol. 198, no. 2020, pp. 512–517, 2021, doi: 10.1016/j.procs.2021.12.278.
- [17] S. Wulandari, Y. I. Mukti, and T. Susanti, "Optimization of the Artificial Neural Network Algorithm

- with Genetic Algorithm in Stroke Prediction," *Sinkron*, vol. 8, no. 2, pp. 1056–1063, 2024, doi: 10.33395/sinkron.v8i2.13609.
- [18] H. Hendriyana, I. M. Karo Karo, and S. Dewi, "Analisis perbandingan Algoritma Support Vector Machine, Naive Bayes dan Regresi Logistik untuk Memprediksi Donor Darah," *J. Teknol. Terpadu*, vol. 8, no. 2, pp. 121–126, 2022, doi: 10.54914/jtt.v8i2.581.
- [19] H. S. W. Hovi, A. Id Hadiana, and F. Rakhmat Umbara, "Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM)," *Informatics Digit. Expert*, vol. 4, no. 1, pp. 40–45, 2022, doi: 10.36423/index.v4i1.895.
- [20] D. P. Sinambela, H. Naparin, M. Zulfadhilah, and N. Hidayah, "Implementasi Algoritma Decision Tree dan Random Forest dalam Prediksi Perdarahan Pascasalin," *J. Inf. dan Teknol.*, vol. 5, no. 3, pp. 58–64, 2023, doi: 10.60083/jidt.v5i3.393.
- [21] D. Y. Utami, E. Nurlelah, and F. N. Hasan, "Comparison of Neural Network Algorithms, Naive Bayes and Logistic Regression to predict diabetes," *J. Informatics Telecommun. Eng.*, vol. 5, no. 1, pp. 53–64, 2021, doi: 10.31289/jite.v5i1.5201.
- [22] Z. Azhari, L. Efrizoni, W. Agustin, and R. Yanti, "Opinion Mining menggunakan Algoritma Deep Learning untuk Menganalisis Penggunaan Aplikasi Jamsostek Mobile," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 666–678, 2023, doi: 10.33022/ijcs.v12i2.3185.
- [23] J. Niimi, "Dynamic Sentiment Analysis with Local Large Language Models using Majority Voting: A Study on Factors Affecting Restaurant Evaluation," 2024, [Online]. Available: <http://arxiv.org/abs/2407.13069>