



**ANALISIS SENTIMEN ULASAN APLIKASI “PINTU
CRYPTO” PADA GOOGLE PLAY STORE
MENGUNAKAN *SUPPORT VECTOR MACHINE***

Nama : Muhamad Sebastian Nugraha



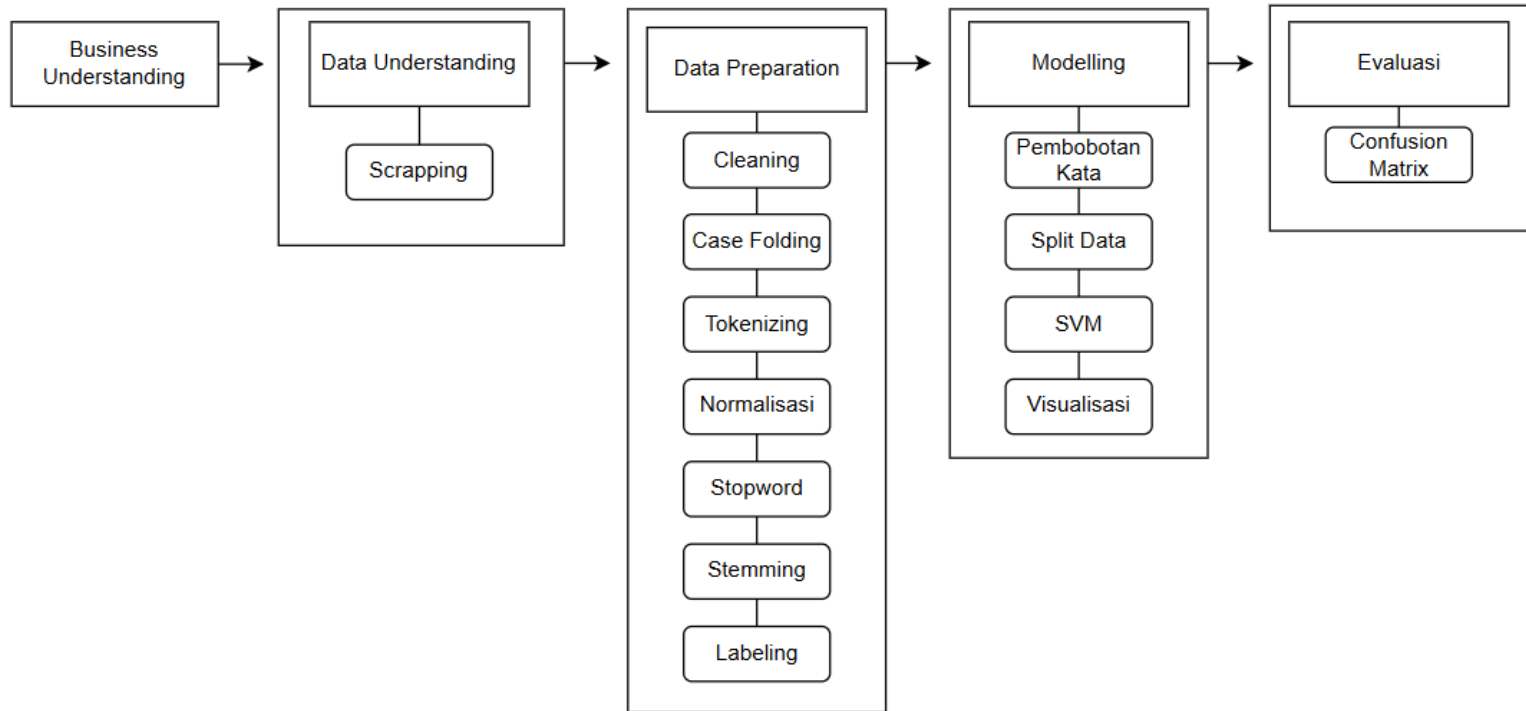
Latar Belakang



- Aplikasi PINTU menunjukkan pertumbuhan pesat sejak diluncurkan, pada di tahun 2021 tercatat telah mencapai 2 juta unduhan, kemudian pada Februari 2025 menjadi lebih dari 9 juta unduhan.
- Jumlah pengguna aktif naik dari 700 ribu (2021) menjadi sekitar 1 juta pada 2024, kemudian diproyeksikan mampu mencapai 28,65 juta investor crypto di Indonesia pada 2025.
- Analisis sentimen ulasan pengguna penting untuk memahami pengalaman pengguna aplikasi serta mengidentifikasi aspek yang perlu ditingkatkan oleh developer.



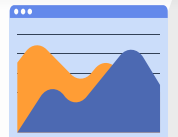
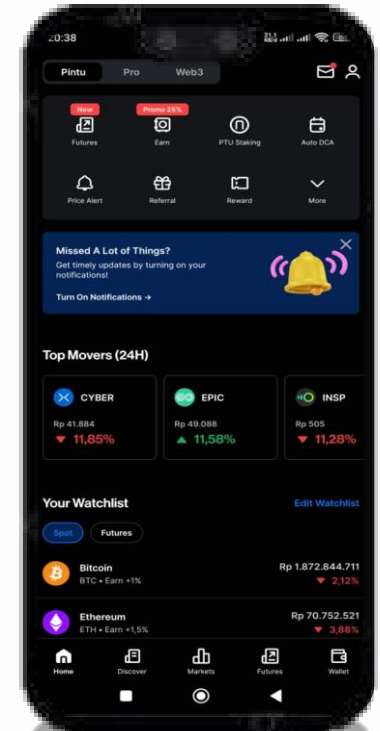
Tahapan CRISP-DM



Business Understanding



- Peran aplikasi Pintu Crypto sebagai platform transaksi aset kripto.
- Aplikasi ini dirancang untuk memfasilitasi pengguna dalam membeli dan menjual berbagai aset kripto seperti Bitcoin, Ethereum, Solana, XRP, dan aset lainnya.
- Fitur yang disediakan mencakup edukasi, peringatan harga, serta keamanan ganda yang bertujuan membangun kepercayaan pengguna dan mempermudah aktivitas jual beli aset kripto.



Data Understanding



reviewId	ID unik dari setiap ulasan untuk membedakan agar tidak terjadi duplikasi data
userName	Nama pengguna yang memberikan ulasan
userImage	Berisi URL gambar profil pengguna
content	Isi ulasan yang ditulis oleh pengguna
score	Nilai rating yang diberikan pengguna terhadap aplikasi, biasanya bintang 1-5
thumbsUpCount	Jumlah pengguna lain yang menyukai ulasan
reviewCreatedVersion	Versi aplikasi ketika ulasan dibuat
at	Tanggal dan waktu saat ulasan diposting
replyContent	Berisi balasan dari pengembang aplikasi terhadap ulasan pengguna
repliedAt	Tanggal dan waktu balasan ulasan yang diposting oleh pengembang aplikasi
appVersion	Versi aplikasi yang digunakan pengguna saat menulis ulasan



Scrapping Data



	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt	appVersion
0	cfbe2d3f-7f15-4d8a-9075-004e055eebcc	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	good job	5	0	3.77.0	2025-07-30 14:57:13	None	NaT	3.77.0
1	e57f678f-cc9d-4b14-b71e-d7f9e532563a	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	oke	5	0	3.77.0	2025-07-30 14:32:31	None	NaT	3.77.0
2	abda5a8c-79af-406a-b9be-61580c10fe40	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	smogga semakin berkembang di kenal banyak oran...	5	0	3.77.0	2025-07-30 14:24:17	None	NaT	3.77.0
3	138750e9-7e33-4a99-9019-2e2a6e977498	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	sangat Membantu	5	0	None	2025-07-30 12:19:13	Hi Teman Pintul Terima kasih atas ulasanya. K...	2025-07-31 12:40:30	None
4	21c07fec-2080-4df8-af2b-70d2dcdb1db7	Pengguna Google	https://play-lh.googleusercontent.com/EGemol2N...	kurang memuaskan	3	0	3.77.0	2025-07-30 09:29:20	Hi Teman Pintul Mohon maaf atas ketidaknyamana...	2025-07-31 12:40:30	3.77.0

Data Preparation



1. Cleaning Data

```
1 df_data[['content', 'cleaned']].head(5)
```

	content	cleaned
0	good job	good job
1	oke	oke
2	smogga semakin berkembang di kenal banyak oran...	smogga semakin berkembang di kenal banyak oran...
3	sangat Membantu	sangat Membantu
4	kurang memuaskan	kurang memuaskan

2. Case Folding

```
1 df_data[['cleaned', 'casefolded']].head(5)
```

	cleaned	casefolded
0	good job	good job
1	oke	oke
2	smogga semakin berkembang di kenal banyak oran...	smogga semakin berkembang di kenal banyak oran...
3	sangat Membantu	sangat membantu
4	kurang memuaskan	kurang memuaskan



Data Preparation



3. Tokenizing

```
1 df_data[['casefolded', 'tokenized']].head(5)
```

	casefolded	tokenized
0	good job	[good, job]
1	oke	[oke]
2	smogga semakin berkembang di kenal banyak oran...	[smogga, semakin, berkembang, di, kenal, banya...
3	sangat membantu	[sangat, membantu]
4	kurang memuaskan	[kurang, memuaskan]

4. Normalisasi

	tokenized	normalization
0	[good, job]	[good, kerja]
1	[oke]	[oke]
2	[smogga, semakin, berkembang, di, kenal, banya...	[semoga, semakin, berkembang, di, kenal, banya...
3	[sangat, membantu]	[sangat, menolong]
4	[kurang, memuaskan]	[tidak cukup, memuaskan]



Data Preparation

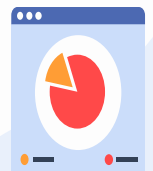


5. Stopword Removal

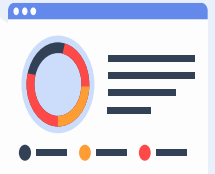
	normalization	stopwords
0	[good, kerja]	[good, kerja]
1	[oke]	[oke]
2	[semoga, semakin, berkembang, di, kenal, banya...]	[semoga, berkembang, kenal, orang, mari, sukse...]
3	[sangat, menolong]	[menolong]
4	[tidak cukup, memuaskan]	[tidak cukup, memuaskan]

6. Stemming

	stopwords	stemming
0	[good, kerja]	good kerja
1	[oke]	oke
2	[semoga, berkembang, kenal, orang, mari, sukse...]	moga kembang kenal orang mari suksesbersama pi...
3	[menolong]	tolong
4	[tidak cukup, memuaskan]	tidak cukup muas



7. Labeling



	stemming	sentiment	polarity
0	good kerja	positive	0.700000
5	salah exchange baik mudah banget mula sukses t...	positive	0.121111
7	gapai mimpi digital pintu cepat lancar mudah aman	positive	0.233333
9	aplikasi pintu gila uninstal aplikasi pintu te...	negative	-0.300000
10	kali coba investasi crypto bitcoin crypto wall...	positive	0.316667

	count
sentiment	
positive	3011
negative	452
dtype: int64	

1. Pembobotan Kata (TF-IDF)



Dokumen 2847:

kadang: 0.7047
nyala: 0.4987
alarm: 0.4579
bagus: 0.1559
ya: 0.1444

Dokumen 2848:

daftar: 0.5216
sukses: 0.5174
fiturnya: 0.4198
lengkap: 0.4027
bagus: 0.2843

Dokumen 2849:

okmhdfauzan: 0.5665
didownload: 0.5201
worth: 0.4525
cocok: 0.2361
paham: 0.2194

2. Split Data

Skenario	Rasio	Data Latih	Data Uji	Total Data
1	80 : 20	2770	693	3463
2	70 : 30	2424	1039	3463
3	60 : 40	2077	1386	3463

3. Model Support Vector Machine

```
[ ] 1 from sklearn.svm import SVC
    2
    3 # Inisialisasi dan training model
    4 svm_model = SVC(kernel='linear')
    5 svm_model.fit(X_train, y_train)
```



SVC

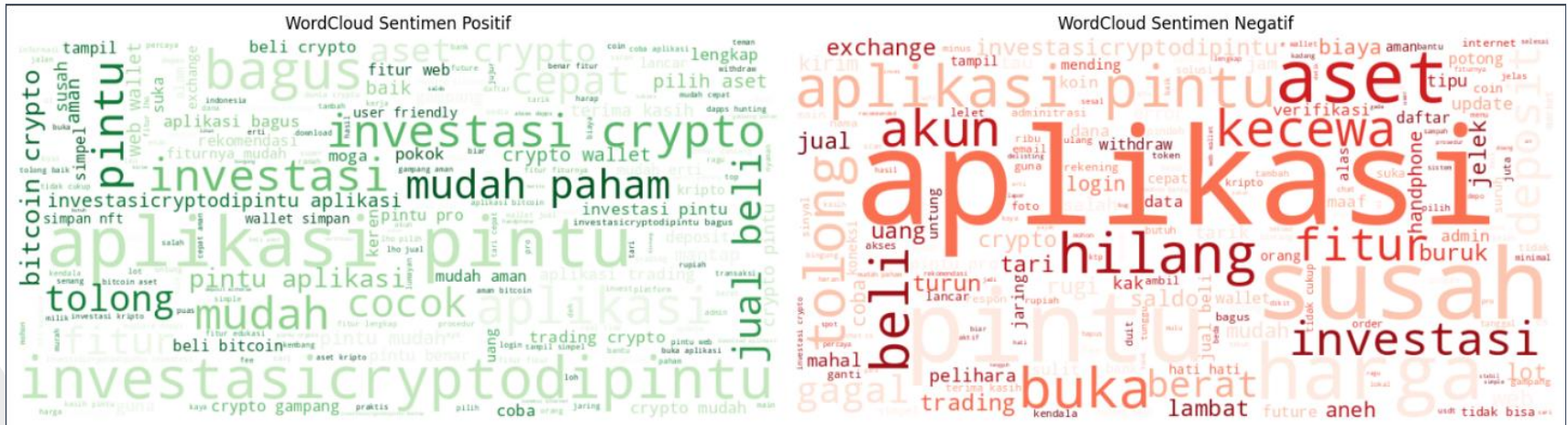


SVC(kernel='linear')

Visualisasi Wordcloud



- Wordcloud sentimen positif didominasi kata seperti **cocok**, **mudah**, **paham**, dan **bagus**, yang menunjukkan kepuasan pengguna terhadap kemudahan aplikasi.
- Sebaliknya, sentimen negatif didominasi kata **susah**, **kecewa**, dan **hilang**, menandakan keluhan terkait akses aplikasi dan kegagalan sistem.



Confusion Matrix



Skenario 1

Akurasi: 0.922077922077922

Laporan Skenario 1:

	precision	recall	f1-score	support
Negatif	0.80	0.58	0.67	95
Positif	0.94	0.98	0.96	598
accuracy			0.92	693
macro avg	0.87	0.78	0.81	693
weighted avg	0.92	0.92	0.92	693

Skenario 2

Akurasi: 0.9162656400384985

Laporan skenario 2:

	precision	recall	f1-score	support
Negatif	0.81	0.57	0.67	153
Positif	0.93	0.98	0.95	886
accuracy			0.92	1039
macro avg	0.87	0.77	0.81	1039
weighted avg	0.91	0.92	0.91	1039

Skenario 3

Akurasi: 0.9148629148629148

Laporan Klasifikasi skenario 3:

	precision	recall	f1-score	support
Negatif	0.77	0.56	0.65	193
Positif	0.93	0.97	0.95	1193
accuracy			0.91	1386
macro avg	0.85	0.77	0.80	1386
weighted avg	0.91	0.91	0.91	1386



Confusion Matrix

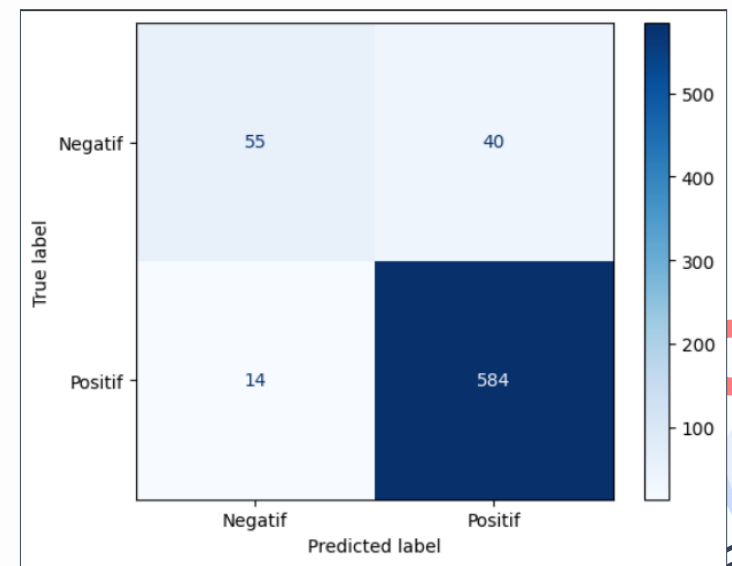


- Hasil pengujian menunjukkan skenario pertama (rasio 80:20) memiliki performa terbaik dengan akurasi **92%**.
- Sebanyak **55 data** dikenali benar sebagai True Negatif, **40 data** masuk False Positif, **584 data** dikenali benar sebagai True Positif, dan **14 data** masuk False Negatif.

Akurasi: 0.922077922077922

Laporan Skenario 1:

	precision	recall	f1-score	support
Negatif	0.80	0.58	0.67	95
Positif	0.94	0.98	0.96	598
accuracy			0.92	693
macro avg	0.87	0.78	0.81	693
weighted avg	0.92	0.92	0.92	693



Kesimpulan dan Saran

Kesimpulan

Dapat dibuat kesimpulan bahwa model Support Vector Machine mampu mendapatkan nilai akurasi sebesar 92%. Program ini dilakukan menggunakan tahap CRISP-DM. Dengan jumlah 5000 data yang dikumpulkan, pelabelan menggunakan textblob, dan split data menggunakan 3 skenario, yaitu 80:20, 70:30, dan 60:40. Analisis sentimen terhadap aplikasi PINTU memberikan hasil yang cukup baik dengan dibuktikan hasil sentimen positif lebih banyak dibandingkan sentimen negatif serta didukung dengan hasil visualisasi wordcloud sehingga memberikan gambaran jelas dari persepsi pengguna.

Saran

Saran yang dapat diberikan adalah dilakukan penyeimbangan jumlah data ulasan positif dan ulasan negatif. Model dapat ditingkatkan dengan melakukan eksperimen pada parameter kernel dan hyperparameter tuning pada Support Vector Machine. Menggunakan metode lain untuk proses labeling seperti BERT untuk mendapatkan hasil kualitas labeling yang lebih baik.



SEKIAN DAN TERIMAKASIH