

Analisis Text Mining pada Sosial Media Twitter Menggunakan Metode Support Vector Machine (SVM) dan Social Network Analysis (SNA)

Tri Putri Lestari¹✉

¹Independent Researcher

triputrilestarii@gmail.com

Abstract

Online loans are growing rapidly in Indonesia in the last two years. This is because the online loan administration requirements are easier compared to bank financial service loans. Online loans are financial services that provide online-based services. Along with the development of online loans, many illegal online loans have sprung up and often commit violations, such as leaking customer personal information and abusing data by carrying out extreme actions such as terrorizing customers who make online loan transactions. This certainly gets a lot of comments from the public, especially on social media twitter. This study aims to conduct a sentiment analysis to see what phenomena are happening among the public regarding online loans. The data used are tweets or retweets from Twitter social media with #pinjamanonline #pinjol. Twitter social media was chosen because an incident can become a phenomenon if it gets a lot of attention from the community, especially on Twitter social media. In this study, using text mining techniques by applying the Support Vector Machine algorithm to classify sentiments on twitter users regarding online loans. This study also looks at the interactions that occur on social media Twitter using social network analysis (SNA). the results of research and testing of the Support Vector Machine method to classify online loans with an Accuracy value level of 86.6%, with a positive precision of 86%, neutral of 1.00% and negative of 87%, positive recall of 90%, neutral 87% and negative of 26 % and positive F1-Score of 88% neutral 42% and negative 86%. Then at the Social Network Analysis stage there is the most influential account, namely influencer @alvinline21 with 1402 nodes.

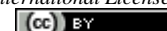
Keywords: Social Network Analysis, Support Vector Machine, Text Mining, Twitter, Online Loan.

Abstrak

Pinjaman Online berkembang dengan pesat di Indonesia dengan rentang waktu dua tahun terakhir. Ini disebabkan karena persyaratan administrasi pinjaman online yang lebih dipermudah di dibandingkan dengan pinjaman layanan keuangan perbankan. Pinjaman online merupakan jasa keuangan yang menyediakan layanan berbasis online. Seiringan dengan perkembangan pinjaman online tersebut, banyak pinjaman online ilegal bermunculan dan sering melakukan pelanggaran, seperti membocorkan informasi pribadi nasabah dan melakukan penyalahgunaan data dengan menjalankan aksi ekstrem seperti melakukan teror kepada nasabah yang melakukan transaksi pinjaman online. Hal ini tentunya mendapat banyak komentar dari kalangan masyarakat khususnya di media sosial twitter. Penelitian ini bertujuan untuk melakukan Analisa sentiment guna melihat fenomena apa yang sedang terjadi di kalangan masyarakat mengenai pinjaman online. Data yang digunakan merupakan tweet atau retweet dari media sosial twitter dengan #pinjamanonline #pinjol. Media sosial twitter di pilih di karenakan suatu kejadian dapat menjadi fenomena jika mendapat banyak perhatian dari kalangan masyarakat terutama di media sosial twitter. Pada penelitian ini menggunakan teknik text mining dengan menerapkan algoritma Support Vector Machine guna melakukan klasifikasi sentimen pada pengguna twitter mengenai pinjaman online. Pada penelitian ini juga dilihat interaksi yang terjadi di sosisl media twitter dengan menggunakan sosial network analysis (SNA). hasil penelitian dan pengujian metode Support Vector Machine untuk mengklasifikasi pinjaman online dengan tingkat nilai Accuracy sebesar 86.6%, dengan precision sebesar positif sebesar 86 %, netral 1.00% dan negatif sebesar 87%, recall positif sebesar 90%, netral 87% dan negatif sebesar 26% dan F1-Score positif sebesar 88% netral 42% dan negatif sebesar 86%. Kemudian pada tahap Social Network Analysis terdapat akun paling berpengaruh yaitu influencer @alvinline21 dengan 1402 nodes.

Kata kunci: Social Network Analysis, Support Vector Machine, Text Mining, Twitter, Pinjaman Online.

INFEB is licensed under a Creative Commons 4.0 International License.



1. Pendahuluan

Pinjaman Online(pinjol) tumbuh dengan pesat di Indonesia dengan rentang waktu 2 tahun terakhir [1]. Ini disebabkan karena persyaratan administrasi pinjol yang lebih dipermudah di dibandingkan dengan pinjaman layanan keuangan perbankan [2]. Kemudahan dan kecepatan yang ditawarkan

memunculkan daya tarik tersendiri. Selain itu, situasi ekonomi yang pelik akibat pandemi Covid-19 menjadikan tingginya permintaan pinjam meminjam di masyarakat. Pinjol merupakan salah satu fasilitas pinjaman uang oleh badan tertentu dengan memanfaatkan teknologi secara online [3]

Semenjak Otoritas Jasa Keuangan (OJK) menerbitkan aturan tentang Peer to Peer lending melalui tata tertib Nomor 77/POJK.01/2016 [4]. Industri jasa keuangan mengalami banyak perkembangan dengan sangat pesat dan memberikan masyarakat pilihan untuk meminimum madu maupun racun [5]. Dengan kemajuan teknologi serta kemudahan yang disediakan sehingga banyak alibi warga memanfaatkan jasa industri keuangan [6]. Namun, bersamaan dengan perkembangan tersebut, banyak pinjol ilegal yang kedapatan sering melakukan pelanggaran, seperti mengungkapkan informasi individu nasabah serta menyalahgunakan informasi pribadi nasabah dengan melakukan tindakan berbahaya seperti melakukan ancaman dan teror terhadap nasabah dalam melaksanakan penagihan pinjaman [7]. Dengan maraknya kasus pinjol ilegal, tidak sedikit masyarakat yang menjadi korban, Sehingga memunculkan banyak sentiment di masyarakat khususnya di sosial media.

Dari fenomena yang terjadi maka penelitian ini akan menganalisis *text mining* pada sosial media twitter terhadap opini masyarakat mengenai pinjol. Dengan adanya penelitian ini nantinya akan melakukan analisis terhadap opini atau pendapat masyarakat mengenai pinjol yang di posting di *twitter* apakah itu bersifat positif, negatif maupun netral. Untuk mewujudkan hal tersebut penelitian ini akan menggunakan sentiment analisis dan *Social Network Analysis* (SNA).

Analisis sentimen ataupun opinion mining adalah proses menguasai, mengekstrak serta mencerna informasi tekstual secara otomatis buat memperoleh data sentimen yang tercantum dalam sesuatu kalimat opini [8]. Analisis sentimen dibuat guna melihat komentar ataupun kecenderungan opini terhadap suatu permasalahan ataupun objek oleh seorang, apakah cenderung beropini negatif ataupun positif [9]. Analisa sentimen juga merupakan salah satu cara untuk mengukur akurasi dalam menentukan sentimen. Banyak metode yang digunakan untuk mengukur akurasi pada *sentimen analysis*, diantaranya *Naïve Bayes Classifier* [10], *K-nearest neighbors algorithm* [11], dan lain sebagainya.

SVM dapat dijelaskan secara sederhana menjadi usaha mencari hyperplane terbaik yang berfungsi menjadi pemisah 2 buah kelas di *input space* [12]. SVM juga dikenal menjadi teknik pembelajaran mesin (*machine learning*) paling terkini setelah pembelajaran mesin sebelumnya yg dikenal menjadi *Neural Network* (NN) [13]. SVM merupakan teknik untuk melakukan klasifikasi opini, Metode SVM dipilih dikarenakan SVM memiliki di nilai hasil yang cukup tinggi pada tingkat keakuratan dibandingkan dengan algoritma yang lain, ini dapat di lihat pada penelitian terdahulu [13], [9], [14]. Selain itu penelitian ini akan menggunakan *Social Network Analysis* (SNA) untuk melihat influencer yang berpengaruh di sosial media twitter. Metode SNA merupakan metode yang mempelajari hubungan antar satu unit entitas dengan

unit entitas lainnya menggunakan bantuan teori *graph* [15].

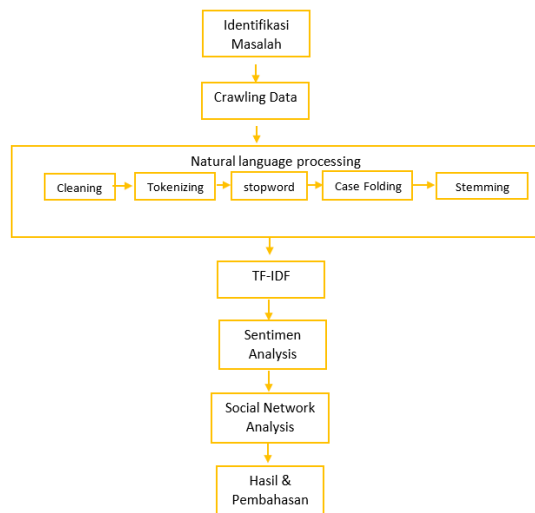
Penelitian yang dilakukan dengan menggunakan Metode *Support Vector Machine* (Svm) guna melihat kebijakan Lockdown yang dilakukan oleh pemerintah jakarta dengan mengambil berjumlah 2000 data di media sosial, hasil penelitian ini dinilai cukup baik di karnakan memiliki akurasi sebesar 74% [16], kemudian penelitian mengenai pembagian terstruktur mengenai Ujaran Kebencian menggunakan menggunakan *Support Vector Machine* di sosial media *twitter* menggunakan akurasi paling tinggi dengan Kernel RBF sebesar 93% penelitian ini memakai data sebesar 700 data *Twitter* [17]. Analisis Sentimen Media Sosial *Twitter* dengan Kasus Pembatasan Sosial Berskala Besar di Jakarta Menggunakan *Algoritma Support Vector Machine* dengan menggunakan data *tweet* sebanyak 1080 opini Proses penjabaran dilakukan menggunakan empat skenario pemisahan data yaitu 60:40, 70:30, 80:20, 90:10 menggunakan akibat rasio pemisahan data 90:10 menggunakan fitur ekstraksi TFIDF lebih unggul dengan nilai akurasi 85,185% dan F1-Score 72,413% [18].

Social Network Analysis (SNA) suatu proses yang mempelajari struktur sosial melalui penggunaan *Network Science* [19]. *Network Science* artinya bidang kajian akademik baru yang menelaah jaringan kompleks mirip jaringan telekomunikasi, jaringan personal personal komputer, jaringan biologis, serta jaringan sosial, kajian ini mendasarkan dirinya di *Graph Theory* [20]. Pendekatan SNA, dimana visualisasi jaringan dimodelkan dengan menggunakan metode *graph* tipe undirected [21]. Penelitian terkait dengan SNA, dilakukan untuk melihat interaksi antar aktor di media sosial salah satunya pada *hastag #pinjaman online, #pinjol*.

Langkah-langkah yang akan dilakukan pada penelitian ini yaitu, Langkah yang melakukan sentimen analisa guna melihat komentar ataupun kritikan *netizen* mengenai Pinjol dengan menggunakan SNA. Langkah selanjutnya penelitian ini melihat hubungan antar *tweet* menggunakan metode SNA. Penelitian ini bertujuan untuk melihat fenomena apa yang sedang terjadi di kalangan masyarakat mengenai pinjaman online, sehingga hasil yang di dapat dari penelitian ini diharapkan dapat berguna bagi masyarakat dalam mempertimbangkan resiko tentang pinjaman online dan meminimalisir adanya penipuan online yang sering terjadi.

2. Metodologi Penelitian

Dalam melakukan sebuah penelitian data dan informasi bersifat objektif yang akan digunakan sebagai titik acuan dalam penelitian, dengan adanya data-data tersebut diharapkan penelitian yang dihasilkan adalah penelitian yang berkualitas. Peroses dalam melakukan penelitian ini digambarkan dalam sebuah flowchart pada Gambar 1.



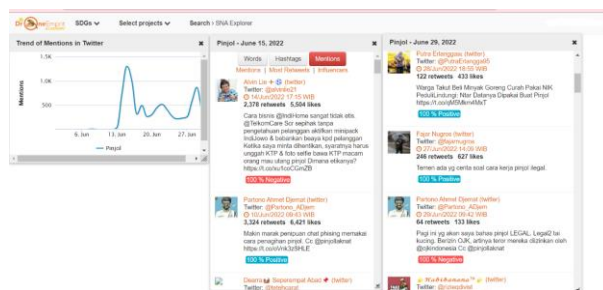
Gambar 1. Metodologi Penelitian

a. Identifikasi Masalah

Identifikasi ini dilakukan sebagai langkah awal penelitian. Jadi, secara ringkas, identifikasi adalah mendefinisikan masalah penelitian. Selain itu, identifikasi masalah juga dapat diartikan sebagai proses dan hasil pengenalan masalah atau inventarisasi masalah.

b. *Crawling Data*

Crawling data merupakan pengumpulan data di mana pengumpulan data dilakukan dengan mengakses portal Drone Emprit Academy (DEA) yang telah disediakan oleh Universitas Islam Indonesia yang merupakan sebuah *system* yang dapat memonitor dan bisa melakukan analisis media *social online*. Data yang diambil berjumlah 5000 tweet. Pencarian data twitter berdasarkan keyword berbahasa Indonesia yaitu “pinjaman online” dan “pinjol”. Berikut adalah tampilan *platform* DEA UI yang di sajikan pada Gambar 2.



Gambar 2. Tampilan Platform DEA UI

c. *Natural language processing* (NLP)

Natural Language Processing (NLP) data pada penelitian dilakukan menggunakan 5 tahap, yaitu:

- Cleaning* yaitu membersihkan *text* dari tanda baca atau HTML tag Bila hasil *crawling* asal *website* [22].

- Case folding* merupakan proses pada *text preprocessing* yang dilakukan buat menyamakan karakter pada data [23].
 - Tokenisasi* adalah proses pemotongan *string input* sesuai istilah yang menyusunnya dan membedakan karakter-karakter eksklusif yang dapat diperlakukan menjadi pemisah kata atau bukan [24].
 - Stopwords*, yaitu menghilangkan kata yang dianggap kurang efektif [25]
 - Stemming* merupakan proses di mana mengubah kata menjadi bentuk dasarnya [26].
- d. Proses SNA pada penelitian ini digunakan untuk melihat pro-kontra yang terjadi sepanjang tahun 2020. Proses SNA dilihat dari portal drone emprit academic (dea.uui.ac.id). Hastag #Pinjamanonline #pinjol Jenis data dalam penelitian ini adalah data kuantitatif dari 01 juni 2022 sampai dengan 30 juni 2022 yang memiliki selang waktu 30 hari Hal ini dilakukan untuk melihat seberapa banyak masyarakat yang beropini mengenai Pinjol di sosial media twitter.

- Proses pembentukan analisis sentimen memakai Metode Support Vector Machine (SVM). setelah dilakukan proses analisis sentimen, maka membuat akurasi dari metode.

f. Seleksi Fitur

Pada seleksi fitur ini di buat menaikkan kinerja *classifier* yg berguna untuk menaikkan akurasi dan mengurangi waktu komputasi. dalam hal ini fitur yang digunakan yaitu TF-IDF [27]. Hasil dari seleksi fitur TF-IDF disajikan pada Tabel 1.

Tabel 1 Data Training

Dokumen	Setelah
K1	lebih horror dr terror pinjol
K2	tambahan, jangan sampe kamu gantung diri gara-gara namamu dipinjem buat pinjol ya, udah kejadian ditempat sya, nama dipinjem temen, hutang gak dibayar, dia yang diuber sampe depresi akhirnya gantung diri
K3	Nonton timnas, beli tiketnya suruh siapin nik sama selfie pake ktp lu federasi bola apa jasa pinjol
K4	Pinjaman, online, ilegal, yang, meresahkan
K5	yuk pak yg butuh pinjaman uang untuk kebutuhan, dg syarat mudah no admin auto acc dan Langsung cair ke rekening klik link dibawah ini ya pak tanpa biaya admin, pinjaman online langsung cair ke rekening hanya modal ktp

g. TF-IDF

Pada tahapan TF-IDF ini, kata-istilah yang tersusun dalam setiap *tweet* untuk diberikan bobot dengan cara perkalian nilai *Term Frequency* (TF) menggunakan *Inverse Document Frequency* (IDF). Hasil dari TF-IDF disajikan pada Tabel 2.

Tabel 2 Perhitungan TF-IDF

Kata	TF					IDF
	K1	K2	K3	K4	K5	
Acc	0	0	0	0	1	$\text{Log}(5/1) = 0.6989$
Admin	0	0	0	0	1	$\text{Log}(5/1) = 0.6989$
Akhimya	0	1	0	0	0	$\text{Log}(5/1) = 0.6989$
Apa	0	0	1	0	0	$\text{Log}(5/1) = 0.6989$
langsung	0	0	0	0	1	$\text{Log}(5/1) = 0.6989$
Biaya	0	0	0	0	1	$\text{Log}(5/1) = 0.6989$
Bola	0	0	1	0	0	$\text{Log}(5/1) = 0.6989$
Buat	0	1	0	0	0	$\text{Log}(5/1) = 0.6989$
Butuh	0	0	0	0	1	$\text{Log}(5/1) = 0.6989$
Cair	0	0	0	0	1	$\text{Log}(5/1) = 0.6989$

h. Hasil dan Pembahasan

Pada tahap ini, akan membahas hasil yang diperoleh dari pengujian yang dilakukan menggunakan beberapa aplikasi.

2.1. Klasifikasi Menggunakan Metode SVM

Proses dengan membuat data *training* dan *testing* dari hasil *crawling* data yang dilakukan dengan menggunakan rasio *dataset* 80% data *training* dan 20% data *testing* disajikan pada Tabel 3.

Tabel 3. Data *Training* dan *Testing*

No	Pembagian data		Jumlah Data	
	Data <i>Training</i>	Data <i>Testing</i>	Data <i>Training</i>	Data <i>Testing</i>
1	80%	20%	4000	1000

Data pelatihan 80% dan data pengujian 20% yang akan diklasifikasikan menggunakan model pembelajaran SVM dengan menggunakan *confusion matrix* yang disajikan pada Tabel 4.

Tabel 4. *Confusion matrix*

Data Aktual	Data Prediksi		
	Negatif	Netral	Positif
Negatif	TNg	NgN	FN
Netral	NNg	TN	NP
Positif	FP	PN	TP

Pada Tabel 4 dijelaskan *Confusion matrix* merupakan hasil prediksi menggunakan mesin klasifikasi SVM yang diukur performa dari tiap-tiap kelas dengan cara menghitung precision, recall, dan F1-score, keterangan:

TNg = Kelas kata negatif terprediksi negatif

NNg = Kelas kata negatif terprediksi netral

FN = Kelas kata negatif terprediksi kata positif

NgN = Kelas kata netral terprediksi kata negatif

TN = Kelas kata terprediksi netral

NP = Kelas kata netral terprediksi sebagai kata positif

FP = Kelas kata positif terprediksi negatif

PN = Kelas kata positif terprediksi netral

TP = Kelas kata prediksi benar bernilai positif

Dari perhitungan manual tersebut sesuai dengan hasil yang diberikat oleh bahasa pemrograman *python* pada percobaan pertama ini dengan proporsi data 80:20 yang disajikan pada Gambar 3

Here is the classification report:				
	precision	recall	f1-score	support
Negative	0.86	0.90	0.88	271
Neutral	1.00	0.27	0.42	15
Positive	0.87	0.86	0.86	214
accuracy			0.87	500
macro avg	0.91	0.68	0.72	500
weighted avg	0.87	0.87	0.86	500

Gambar 3. Hasil SVM dengan proporsi data 80:20

2.2 Analisis Sentimen

Pada pembahasan analisis sentimen, menggunakan metode pengujian, dan data pada penelitian ini dibagi menjadi proporsi data, yakni 80:20 dengan *confusion matrix* yang disajikan pada Tabel 4.

Tabel 4. Hasil 80:20 SVM

Data Aktual	Data Prediksi			Precision	Recall	F1-Score
	Neg	Net	Pos			
Neg	150	3	20	0.86	0.90	0.88
Net	11	19	26	1.00	0.87	0.42
Pos	17	4	144	0.87	0.26	0.86

Nilai akurasi Support Vector Machine

Akurasi = 86.6%

Presisi = $(TP)/(TP+FP)$

$$= (97)/(97+94) \times 100\% = 50,7\%$$

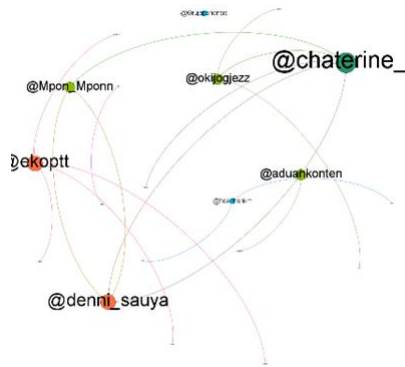
Hasil analisis dan pengujian yang telah dilakukan menggunakan data uji sebanyak 4000 komentar dari masyarakat mengenai topik pinjaman online dan pinjol menggunakan media massa twitter pada metode SVM menghasilkan nilai Akurasi sebesar 86.6% dan presisi 50,7%.

2.3. Perhitungan Nilai Centrality

Pada tahap merupakan tahap perhitungan nilai centrality (*degree centrality*, *closeness centrality*, dan *betweenness centrality*) node atau aktor untuk mengidentifikasi aktor berpengaruh dengan jumlah interaksi yang tinggi.

Perhitungan degree centrality pada hastag pinjaman online dan pinjol dengan node @denni_sauya di mana perhitungan degree centrality yang dilakukan dengan bantuan software gephi di dapat 4 (Empat) interaksi yang terjadi, maka dilakukan perhitungan menggunakan Rumus (1).

$$C_D(n_i) = d(n_i) \quad (1)$$



Gambar 4. Interaksi Degree Centrality

Pada Gambar 4 pada node @denni_sauya dapat dilihat degree centrality terdapat ada 4 (Empat) interaksi node. Maka di ketahui:

$$dn_{i=4}$$

3. Hasil dan Pembahasan

3.1 Support Vector Machine

Pengujian yang dilakukan yaitu dengan fitur TF-IDF bertujuan untuk mengetahui seberapa besar nilai *Accuracy*, *Precision*, *recall* dan *F1-Score*, berdasarkan pengujian memiliki hasil yang ditunjukkan pada tabel 4 Ini merupakan hasil *dataset* dengan perbandingan 80:20. Berikut hasil dari proses percobaan menggunakan *Python* dalam klasifikasi SVM yang disajikan dalam Gambar 5.

Here is the classification report:				
	precision	recall	f1-score	support
Negative	0.86	0.90	0.88	271
Neutral	1.00	0.27	0.42	15
Positive	0.87	0.86	0.86	214
accuracy			0.87	500
macro avg			0.72	500
weighted avg			0.86	500

Gambar 5. Hasil Klasifikasi SVM

Hasil analisis dan pengujian yang telah dilakukan menggunakan data uji sebanyak 4000 tweet mengenai topik pinjaman online dan pinjol menggunakan media massa twitter pada metode SVM menghasilkan nilai presisi 91 %, *recall* sebesar 68 % kemudian untuk nilai akurasi menghasilkan nilai persentase sebesar 87 %. Dapat dilihat pada Tabel 5 Rekap hasil SVM

Tabel 5. Rekap hasil SVM

	Support Vector Machine			
	Accuracy	Recall	Preccision	F1-Score(%)
Positif	86.6 %	0.90 %	0.86 %	0.88
Netral	86.6 %	0.87 %	1.00 %	0.42
Negatif	86.6 %	0.26 %	0.87 %	0.86

Dari proporsi data percobaan yang dilakukan menggunakan SVM pada Tabel 5 di sajikan data yang merupakan rekap dari hasil percobaan menggunakan SVM yang terdapat pada gambar 4 Hasil Klasifikasi SVM.

3.1.1 Word Cloud

Word Cloud menggambarkan kata-kata yang paling sering muncul dan menjadi topik utama dari 5000 tweets yang terdapat pada postingan di media sosial twitter tersebut. Word cloud merupakan topik teratas digambarkan melalui kumpulan kata. Kata yang paling sering muncul akan ditampilkan dalam ukuran yang lebih besar dari yang lainnya, seperti yang telah di sajikan pada Gambar 6 dan 7 visualisasi berbentuk wordcloud dengan 2 kelas, yaitu positif, dan negatif.

visualisasi wordcloud untuk sentiment “Positive” :



Gambar 6. *Wordcloud* dengan label *positive*

Berikut ini visualsiasi worcloud untuk sentiment “Negative” :



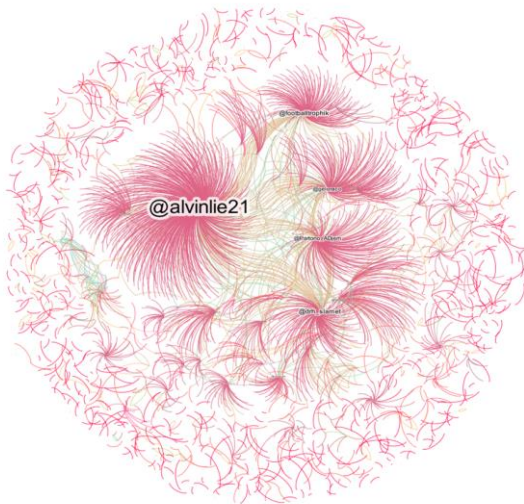
Gambar 7. Wordcloud dengan label negative

3.2 Social Network Analysis (SNA)

Pada SNA yang terlihat pada gambar 3, terdapat 3 *influencer* yang begitu dominan, yaitu @alvinline21, @drh_slamet dan @Footballtrophlk. Ketiga influencer ini merupakan orang yang berpengaruh di media sosial *twitter*. Sehingga pemerintah dapat menjadikan ketiga

akun tersebut sebagai influencer terhadap masyarakat dalam memberikan edukasi mengenai pinjaman online.

Perhitungan hasil data *hashtag* pinjaman online dan pinjol di media sosial *twitter* menggunakan *software gephi* versi 0.9.6 terhadap nilai *degree centrality*, *closeness centrality*, dan *betweenness centrality* maka di ketahui aktor yang berpengaruh terhadap interaksi jejaring sosial adalah @alvinlie21 yang unggul pada nilai *degree centrality*, *closeness centrality*, dan *betweenness*, akun @drh_slamet, dengan nilai *degree* 208, @Footballtrophik dengan nilai *degree* 207, dan di susul oleh @Partono_ADjem dengan nilai *degree* 180. Adapun gambar model jaringan dapat di lihat pada Gambar 8.



Gambar 8. Visualisasi jaringan #pinjol #pinjamanonline

SNA ialah hubungan yang terjadi terhadap user yang dilambangkan menggunakan titik (nodes) serta hubungan antar user tersebut yg dilambangkan dengan garis (edges) [28]. SNA membantu untuk memahami korelasi sosial yang melambangkan user dengan titik (nodes) dan korelasi antar user dilambangkan menggunakan garis (edges) di *Online Social Network* (OSN) [26]. Pada Tabel 8 merupakan 10 tertinggi dari nodes SNA influencer.

Tabel 8. Nodes SNA Influencer

Name	Type	Size
@alvinlie21	person	1402
@drh_slamet	person	1385
@Partono_ADjem	person	315
@geloraco	person	307
@footballtrophik	person	197
@DivHumas_Polri	person	185
@BossTemlen	person	174
@ltrenggalek	person	145
@quintta1217	person	144
@HumasPoldaJatim	person	139

3.2.1 Word Cloud

Word Cloud menggambarkan katakata yang paling sering muncul dan menjadi topik utama dari 5000 tweets yang terdapat pada postingan di media sosial *twitter* tersebut. *Word Cloud* merupakan topik teratas digambarkan melalui kumpulan kata. Kata yang paling sering muncul akan ditampilkan dalam ukuran yang lebih besar dari yang lainnya, yang disajikan pada Gambar 9 *Word Cloud* pinjaman online:



Gambar 9. Word Cloud pinjaman online

4. Kesimpulan

Berdasarkan hasil penelitian dan pengujian metode *Support Vector Machine* untuk mengklasifikasi pinjaman *online* dengan *tweet* berbahasa Indonesia menggunakan ekstraksi fitur TF-IDF didapat kesimpulan bahwa penggunaan metode *Support Vector Machine* tingkat nilai *Accuracy* sebesar 86.6%, dengan *Precision* sebesar *Positif* sebesar 86 %, *Netral* 1.00% dan *Negatif* sebesar 87%, *Recall Positif* sebesar 90%, *Netral* 87% dan *Negatif* sebesar 26% dan *F1-Score Positif* sebesar 88% *Netral* 42% dan *Negatif* sebesar 86%. Setelah menyelesaikan serangkaian tahapan terhadap analisa data sentimen pada *twitter* dengan objek Pinjaman *Online*. Kemudian pada tahap SNA terdapat 3 akun paling berpengaruh yang terdapat pada 2 cluster. Cluster pertama terdapat akun influencer @alvinlie21 dan @drh_slamet. Kemudian pada cluster ke 2 terdapat @Footballtrophik yang paling berpengaruh pada *hashtag* #pinjaman online dan #pinjol.

Daftar Rujukan

- [1] Akyuwen, R., Nanere, M., & Ratten, V. (2022). Technology entrepreneurship: Fintech lending in Indonesia. In *Entrepreneurial innovation* (pp. 151-176). Springer, Singapore. DOI: https://doi.org/10.1007/978-981-16-4795-6_14
- [2] Syarvina, W., & Sudiarti, S. (2022). Analisa Risiko Pinjaman Online Ilegal Dalam Praktik Teknologi Finansial. *Jurnal Riset Akuntansi dan Bisnis*, 22(1), 18-28. DOI: <http://dx.doi.org/10.30596%2F8939>
- [3] Wahyuni, R. A. E., & Turisno, B. E. (2019). Praktik Finansial Teknologi Ilegal Dalam Bentuk Pinjaman Online Ditinjau Dari Etika Bisnis. *Jurnal Pembangunan Hukum Indonesia*, 1(3), 379-391. DOI: <https://doi.org/10.14710/jphi.v1i3.379-391>
- [4] Sopiyan, R. (2022). Peraturan Otoritas Jasa Keuangan (Ojk) Terhadap Perlindungan Konsumen Atas Data Pribadi Dalam

- Transaksi Peer To Peer Lending (P2p Lending) Berbasis Teknologi Informasi (Doctoral Dissertation, Universitas Siliwangi. DOI: <https://doi.org/10.31289/jiph.v8i1.512>
- [5] Marsden, G., & Docherty, I. (2021). Mega-disruptions and policy change: Lessons from the mobility sector in response to the Covid-19 pandemic in the UK. *Transport Policy*, 110, 86-97. DOI: <https://doi.org/10.1016/j.tranpol.2021.05.015>
- [6] Levidow, L. (2022). Green New Deals: What Shapes Green and Deal Capitalism Nature Socialism, 1-22. DOI: <https://doi.org/10.1080/10455752.2022.2062675>
- [7] Sidiq, V. A. R. A., & Setiawan, H. (2022). Analisis Framing Pemberitaan Kasus Pinjaman Online Warga Negara China pada Media Online CNNIndonesia. com dan Nasional Tempo. com. Edukatif: *Jurnal Ilmu Pendidikan*, 4(1), 851-861. DOI: <https://doi.org/10.31004/edukatif.v4i1.1935>
- [8] Bourequat, W., & Mourad, H. (2021). Sentiment analysis approach for analyzing iPhone release using support vector machine. *International Journal of Advances in Data and Information Systems*, 2(1), 36-44. DOI: <https://doi.org/10.25008/ijadis.v2i1.1216>
- [9] Fitriyah, N., Warsito, B., & Di Asih, I. M. (2020). Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (SVM). *Jurnal Gaussian*, 9(3), 376-390. DOI: <https://doi.org/10.14710/j.gauss.v9i3.28932>
- [10] Salsabila, S. M., Murtopo, A. A., & Fadhilah, N. (2022). Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier. *Jurnal Minfo Polgan*, 11(2), 30-35. DOI: <https://doi.org/10.33395/jmpv11i2.11640>
- [11] Gunawan, R., Septiadi, R., Wenando, F. A., & Mukhtar, H. (2022). K-Nearest Neighbor (KNN) untuk Menganalisis Sentimen terhadap Kebijakan Merdeka Belajar Kampus Merdeka pada Komentar Twitter. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(2), 152-158. DOI: <https://doi.org/10.37859/coscitech.v3i2.3841>
- [12] Denny, Y. R., Permata, E., & Assaat, L. D. (2022). Classification of diseases of banana plant fusarium wilted banana leaf using support vector machine. Gravity: *Jurnal Ilmiah Penelitian dan Pembelajaran Fisika*, 8(1). DOI: <http://dx.doi.org/10.30870/gravity.v8i1.15893>
- [13] Alhaq, Z., Mustopa, A., Mulyatun, S., & Santoso, J. D. (2021). Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter. *Journal Of Information System Management (Joism)*, 3(1), 16-21. DOI: <https://doi.org/10.24076/joism.2021v3i2.558>
- [14] Mutawalli, L., Zaen, M. T. A., & Bagye, W. (2019). Klasifikasi Teks Sosial Media Twitter Menggunakan Support Vector Machine (Studi Kasus Penusukan Wiranto). *Jurnal Informatika dan Rekayasa Elektronik*, 2(2), 43-51. DOI: <https://doi.org/10.36595/jire.v2i2.117>
- [15] Anam, M. K., Lestari, T. P., Firdaus, M. B., & Fadli, S. (2021). Analisis Kesiapan Masyarakat Pada Penerapan Smart City di Sosial Media Menggunakan SNA. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 69-81. DOI: <https://doi.org/10.29207/resti.v5i1.2742>
- [16] Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), 31-37. DOI: <https://doi.org/10.33365/jdmsi.v2i1.1021>
- [17] Rahman, O. H., Abdillah, G., & Komarudin, A. (2021). Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1), 17-23. DOI: <https://doi.org/10.29207/resti.v5i1.2700>
- [18] Saragih, P. S., Winarsyah, D., Hamami, F., & Machado, J. M. (2021, October). Sentiment Analysis of Social Media Twitter with Case of Large Scale Social Restriction in Jakarta using Support Vector Machine Algorithm. In *2021 International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS)* (pp. 1-6). IEEE. DOI: <https://doi.org/10.1109/ICADEIS52521.2021.9701961>
- [19] Powell, K. R., Popescu, M., & Alexander, G. L. (2021). Examining Social Networks in Text Messages About Nursing Home Resident Health Status. *Journal of Gerontological Nursing*, 47(7), 16-22. DOI: <https://doi.org/10.3928/00989134-20210604-02>
- [20] Taira, K., & Nair, A. G. (2022). Network-based analysis of fluid flows: Progress and outlook. *Progress in Aerospace Sciences*, 131, 100823. DOI: <https://doi.org/10.1016/j.paerosci.2022.100823>
- [21] Oroh, A. J., Bandung, Y., & Zagi, L. M. (2021, April). Detection of the Key Actor of Issues Spreading Based on Social Network Analysis in Twitter Social Media. In *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)* (pp. 206-212). IEEE. DOI: <https://doi.org/10.1109/APWiMob51111.2021.9435268>
- [22] Lowphansirikul, L., Polpanumas, C., Rutherford, A. T., & Nutanong, S. (2022). A large english-thai parallel corpus from the web and machine-generated text. *Language Resources and Evaluation*, 56(2), 477-499. DOI: <https://doi.org/10.1007/s10579-021-09536-6>
- [23] Yunanda, G., Nurjanah, D., & Meliana, S. (2022). Recommendation System from Microsoft News Data using TF-IDF and Cosine Similarity Methods. *Building of Informatics, Technology and Science (BITS)*, 4(1), 277-284. DOI: <https://doi.org/10.47065/bits.v4i1.1670>
- [24] Hacquard, A., & Verna, D. (2021, May). A Corpus Processing and Analysis Pipeline for Quickref. In *14th European Lisp Symposium*. DOI: <https://doi.org/10.5281/zenodo.4714443>
- [25] Kochhar, T. S., & Goyal, G. (2022). Design and Implementation of Stop Words Removal Method for Punjabi Language Using Finite Automata. In *Advances in Data Computing, Communication and Security* (pp. 89-98). Springer, Singapore. DOI: https://doi.org/10.1007/978-981-16-8403-6_8
- [26] Hutajulu, T. A., Priyadi, Y., & Gandhi, A. (2022, June). Text Data Processing in Requirement Specifications as a Reference for Similarities Between Use Case Diagrams and Use Case Descriptions for Smart Sleeping Lamp Application Documents. In *2022 IEEE World AI IoT Congress (AIIoT)* (pp. 665-671). IEEE. DOI: <https://doi.org/10.1109/AIIoT54504.2022.9817197>
- [27] Septianingrum, F., Jaman, J. H., & Enri, U. (2021). Analisis Sentimen Pada Isu Vaksin Covid-19 di Indonesia dengan Metode Naive Bayes Classifier. *Jurnal Media Informatika Budidarma*, 5(4), 1431-1437. DOI: <http://dx.doi.org/10.30865/mib.v5i4.3260>
- [28] Prabowo, N. A., Pujiarto, B., Wijaya, F. S., Gita, L., & Alfandy, D. (2021). Social network analysis for user interaction analysis on social media regarding e-commerce business. *International Journal of Informatics and Information Systems*, 4(2), 95-102. DOI: <https://doi.org/10.47738/ijis>