

ANALISIS SENTIMEN FINANCIAL TECHNOLOGY PEER TO PEER LENDING PADA APLIKASI KOINWORKS

Rousyati^{*1}, Windu Gata², Dany Pratmanto³, Nia Kusuma Wardhani⁴

^{1,2}STMIK Nusa Mandiri, Jakarta Pusat, ³Universitas Bina Sarana Informatika, Depok, ⁴Universitas Mercu Buana, Jakarta

Email: ¹ 14002240@nusamandiri.ac.id, ² windu@nusamandiri.ac.id, ³dany.dto@bsi.ac.id,
⁴nia.kusuma@mercubuana.ac.id

*Penulis Korespondensi

(Naskah masuk: 25 November 2020, diterima untuk diterbitkan: 14 Desember 2022)

Abstrak

Bertambahnya jumlah perusahaan *financial technology* (*fintech*) yang terdaftar di Otoritas Jasa Keuangan mengartikan bahwa industri ini semakin dilirik karena dibutuhkan dalam sistem perekonomian di Indonesia. Namun perkembangan *Fintech* P2PL telah menimbulkan beberapa risiko. Pertama, ada risiko gagal bayar, karena tidak ada jaminan atau persyaratan kontak fisik. Kedua, ada risiko yang terkait dengan keamanan data (risiko cyber), tata kelola, dan privasi pelanggan dan juga karena kerentanan sistem dan penyalahgunaan data, baik sengaja atau tidak sengaja. Ulasan yang terdapat pada kolom komentar Google Play dapat dimanfaatkan sebagai sumber data yang dapat diolah dengan data mining. Penelitian ini akan menganalisis mengenai permasalahan yang berkaitan dengan beberapa ulasan tentang *Fintech* P2PL aplikasi Koinworks pada ulasan di Google Play Store serta menentukan hasil akurasi analisis sentimen yang dihasilkan algoritma *Decision Tree*, *K-Nearest Neighbor* dan *Support Vector Machine*. Adapun manfaat dari penelitian ini adalah untuk membantu manajemen aplikasi Koinworks mengenai opini positif atau negatif dari pengguna aplikasi serta dapat memberikan bukti secara empiris untuk teori yang berkaitan sehingga dapat dijadikan sumbangan pemikiran untuk pengembangan teori berikutnya. Algoritma SVM dengan Cross Validation + Parameter Optimization menghasilkan *Accuracy* 91,03% *precision* tertinggi yaitu dengan 96,73% , *recall* 85,34% dan AUC tertinggi yaitu 0,986 yang termasuk dalam *excellent classification*.

Kata kunci: Analisis sentimen, *fintech peer to peer lending*, text mining, DT, KNN, SVM

SENTIMENT ANALYSIS OF FINANCIAL TECHNOLOGY PEER TO PEER LENDING ON KOINWORKS APPLICATION

Abstract

The increasing number of financial technology (*fintech*) companies registered with the Financial Services Authority means that this industry is increasingly being looked at because it is needed in the economic system in Indonesia. However, the development of *Fintech* P2PL has created several risks. First, there is a risk of default, because there are no guarantees or physical contact requirements. Second, there are risks associated with data security (cyber risk), governance, and customer privacy and also because of system vulnerabilities and data abuse, whether intentionally or unintentionally. Reviews contained in the Google Play comments column can be used as a data source that can be shared with data mining. This research will analyze the problems related to some reviews about the *Fintech* P2PL Koinworks application on reviews on the Google Play Store and determine the results of the accuracy of sentiment analysis produced by the *Decision Tree* algorithm, *K-Nearest Neighbor* and *Support Vector Machine*. The benefits of this research are to help the management of Koinworks applications regarding positive or negative opinions of application users and can provide empirical evidence for related theories so that they can be contributed to the development of subsequent theories. SVM algorithm with Cross Validation + Parameter Optimization produces *Accuracy* 91.03% of the highest precision with 96.73 %%, 85.34% recall and the highest AUC of 0.986 which is included in excellent classification.

Keywords: sentiment analysis, *fintech peer to peer lending*, text mining, DT, KNN, SVM

1. PENDAHULUAN

Bertambahnya jumlah perusahaan *financial technology* (*fintech*) yang terdaftar di Otoritas Jasa

Keuangan mengartikan bahwa industri ini semakin dilirik karena dibutuhkan dalam sistem perekonomian di Indonesia. Salah satu *fintech* yang

berpengaruh dalam sistem perekonomian adalah *fintech Peer to Peer lending* (P2PL). Terbukti pada bulan Maret 2020 terdapat 161 *fintech* P2PL yang tercatat di Otoritas Jasa Keuangan (OJK), diantaranya 25 dengan status berizin dan 96 lainnya masih dalam status terdaftar (OJK, 2020).

Peer to Peer Lending merupakan platform yang mempertemukan peminjam dengan pemberi pinjaman secara digital yang unggul dalam menjalankan fungsi *interface* melalui pendanaan selain itu teknologi ini fleksibel dan dapat mengalokasikan dana kepada siapa saja dalam jumlah yang disepakati secara transparan. Layanan *Peer to Peer Lending* dapat menjembati UMKM peminjam yang *credit worthy* menjadi *bankable* dengan menyediakan pinjaman tanpa agunan (Kurniawan, Wardani and Widhayati, 2019).

Penawaran layanan *Fintech P2PL* berbagai fasilitas untuk meminjam dengan persyaratan yang lebih sederhana dari pada bank. Tidak ada persyaratan agunan, pemrosesan lebih cepat dan tidak perlu fisik kehadiran (akses internet adalah semua yang diperlukan untuk mencapai layanan). Namun perkembangan *Fintech P2PL* telah menimbulkan beberapa risiko. Pertama, ada risiko gagal bayar, karena tidak ada jaminan atau persyaratan kontak fisik. Kedua, ada risiko yang terkait dengan keamanan data (risiko *cyber*), tata kelola, dan privasi pelanggan dan juga karena kerentanan sistem dan penyalahgunaan data, baik sengaja atau tidak sengaja (Narain, 2017).

Selain itu, pengawasan OJK layanan *Fintech P2PL* relatif lemah, karena menghadirkan beberapa masalah perlindungan pelanggan. Salah satunya masalahnya adalah metode penagihan yang tidak etis. Penagihan hutang juga dilakukan dengan cara yang tidak pantas, seperti dengan intimidasi, ancaman dan penghinaan. Hal tersebut mendorong pembuatan petisi online yang telah ditandatangani oleh lebih dari 5.000 orang (Change.org, 2018). Praktik buruk ini melanggar Peraturan OJK Nomor 1 Tahun 2013, tentang perlindungan konsumen dari sektor jasa keuangan, dan Menteri Peraturan Komunikasi dan Informasi Nomor 20 Tahun 2016, tentang perlindungan data pribadi dalam sistem elektronik.

Google Play menyediakan data yang berisi Google Play menyediakan data yang berisi layanan tentang ulasan pengguna aplikasi. Sejauh ini, adalah aplikasi terbesar *platform* di dunia. Pada Maret 2018, Google Play telah diunduh lebih dari 3,3 juta kali, sedangkan iOS App Store telah diunduh sekitar 2,03 juta kali (Ignatow and Mihalcea, 2015). Aplikasi ini menyediakan distribusi perangkat lunak sehingga pengguna dapat memperoleh aplikasi langsung dari pengembang selain itu pengguna dapat memberikan ulasan mengenai aplikasi di kolom komentar yang disediakan terdiri dari berbagai macam pendapat baik negatif maupun positif.

Ulasan yang terdapat pada kolom komentar Google Play dapat dimanfaatkan sebagai sumber data

yang dapat di olah dengan *data mining*. *Data mining* adalah metode pencarian pengetahuan dalam basis data untuk menemukan pengetahuan yang bermanfaat dari data. Proses ini untuk mendapatkan desain yang menarik dan hubungan dan dapat dilayani dalam volume besaran data. Dengan demikian *data mining* dapat digunakan sebagai sentimen analisis karena memiliki sejumlah besar (Priyadharsini.C and Thanamani, 2014).

Dalam analisis sentimen klasifikasi dibagi menjadi tiga tingkatan, yaitu tingkat dokumen, tingkat kalimat, dan tingkat aspek. Karena yang digunakan data tekstual maka terdapat kebutuhan untuk menganalisa konsep mengekspresikan sentimen dan menghitung wawasan untuk menjelajahi bisnis. Analisis sentimen merupakan suatu proses untuk menentukan apakah ulasan yang diungkapkan cenderung ulasan positif, negatif atau netral. Opini *mining* tidak memperhatikan topik dari teks tersebut tetapi lebih fokus kepada ekspresi yang digambarkan dari teks opini tersebut.

Menurut *Bisnis Indonesia Fintech Award 2017* KoinWorks merupakan *The Most Innovative Fintech of The Year* yang menjadi salah satu pelopor *platform* investasi online *Peer to Peer Lending application*. KoinWorks dapat diklasifikasikan sebagai perdagangan elektronik karena fitur yang menyatukan pembeli, yaitu investor dan penjual atau peminjam dana dengan sistem online (Nugraha et al., 2019). Dipilihnya aplikasi Koinworks karena menjadi salah satu pelopor investasi *Peer to Peer Lending Fintech* di Indonesia sehingga banyak Usaha Mikro Kecil dan Menengah (UMKM) berhasil memperoleh kredit untuk pengembangan usahanya melalui aplikasi ini (Wulandari, 2017).

Penelitian untuk menganalisis sentimen tentang *review* atau ulasan pengguna perlu dilakukan untuk mengetahui dan memprediksi tingkat kepuasan pengguna. Dalam mengklasifikasi ulasan dapat dilihat dari *rating* yang diberikan pada kolom komentar di google play store, namun hal tersebut kurang akurat sehingga diperlukan analisis sentimen dengan metode klasifikasi data mining.

Penelitian dengan judul Pengaruh Keberterimaan Layanan Peer To Peer Lending Kepada UMKM Sebagai Pengguna Dengan Menggunakan Metode *Technology Acceptance Model* (TAM) dilakukan untuk mengetahui pengaruh *Perceived Ease of Use*, Persepsi Khasiat, Kepercayaan, dan sikap pengguna terhadap penerimaan sistem pinjaman *peer to peer*. (Kurniawan, Wardani and Widhayati, 2019).

Penelitian dengan judul *The Sentiment Analysis of Fintech Users Using Support Vector Machine and Particle Swarm Optimization Method* dilakukan untuk menganalisis sentimen pengguna Fintech di Kota Tegal, terutama aplikasi Ovo karena saat ini memiliki pasar dan promosi yang sangat m

asif. Pengguna yang memilih aplikasi *Fintech* umumnya mempertimbangkan kemudahan,

keamanan, kesesuaian transaksi, kenyamanan, dan cashback. Masalahnya adalah bahwa kepercayaan pengguna pada Fintech masih diragukan. Model yang dihasilkan mendapatkan akurasi tes pertama hasil dilakukan dengan menggunakan Support Vector Machine dengan nilai akurasi: 75,33% +/- 7,77% (mikro: 75,33%) yaitu kemudian diuji lagi dengan tes kedua dengan menggabungkan Support Particle Swarm Optimization (PSO) Berbasis Support Vector Machine dengan nilai terbaik dengan akurasi 82,33% +/- 7,00% (mikro: 82,33%). Maka dapat disimpulkan bahwa pengujian analisis sentimen menggunakan Support Vector Machine akan menjadi lebih baik dengan menggabungkan Support Vector Machine-Based Penguoptimalan Particle Swarm (PSO) (Aji et al., 2019).

Analisis sentimen dan emosi memberikan cara cepat dan mudah untuk menyimpulkan persepsi pengguna mengenai produk, layanan, topik, dan peristiwa, dan dengan demikian menjadikannya berguna bagi bisnis dan badan pemerintah untuk pengambilan keputusan yang efektif. (Balakrishnan, Selvanayagam and Yin, 2020).

Penelitian terkait dengan judul Analisis Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based Dan Support Vector Machine, Go-Pay merupakan bagian dari aplikasi Gojek dan salah satu *fintech* yang paling diminati di Indonesia. (Mahendrajaya, Buntoro and Setyawan, 2019).

Pengkategorian ulasan tersebut secara otomatis, apakah termasuk positif atau negatif. Algoritma yang digunakan dalam penelitian ini adalah Naive Bayes Classifier (NB), dengan optimasi penggunaan Feature Selection (FS) Particle Swarm Optimization. Hasil dari cross validation NB tanpa FS adalah 82.30 % untuk accuracy dan 0.780 untuk AUC (Aaputra et al., 2019).

Penelitian dengan judul *Instagram Sentiment Analysis with Naive Bayes and KNN: Exploring Customer Satisfaction of Digital Payment Services in Indonesia* dilakukan untuk mengetahui kepuasan pelanggan layanan pembayaran digital (Go-Pay, Ovo, dan LinkAja) di Indonesia, dianalisis dari hasil analisis sentimen Instagram, di mana bentuk data tekstual yang digunakan adalah komentar Instagram. Algoritma klasifikasi paling akurat yang diterapkan dalam penelitian ini adalah algoritma classifier KNN, hingga 90%. (Sudira, Diar and Ruldeviyani, 2019).

Penelitian *Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naive Bayes* analisis sentimen dan penambangan pendapat dilakukan untuk melihat kepuasan publik menuju layanan pembayaran digital di Indonesia (OVO, GO-PAY dan LinkAja). Penelitian ini menggunakan algoritma klasifikasi Naive Bayes dan KNN, hasilnya menunjukkan Naive Bayes memiliki akurasi data terbaik saat melakukan validasi silang 20 kali lipat. Akurasi terbaik GO-PAY (83,50%) dicapai setelah berjalan 20 kali lipat dengan algoritma KNN. Untuk

OVO dan LinkAja, akurasi mencapai yang terbaik (84,00% dan 91,00%) setelah cross 20 kali lipat validasi dengan algoritma KNN.. (Wisnu, Afif and Ruldevyani, 2020).

Penelitian *Comparison Analysis Of Social Influence Marketing For Mobile Payment Using Support Vector Machine* Ada banyak layanan keuangan berbasis digital saat ini, salah satunya adalah layanan pembayaran mobile. Pengguna dapat menyimpan uang dan melakukan transaksi online dengan ponsel cerdas mereka melalui aplikasi seluler. Lima penyedia layanan pembayaran seluler dengan pengguna terbanyak di Indonesia, menurut Dailysocial adalah GOPAY, OVO, LinkAja, KOINWORKS, dan PayTren. Penelitian ini menggunakan analisis sentimen untuk mengklasifikasikan opini pengguna ke dalam kelas positif dan negatif. Metode klasifikasi yang digunakan adalah Support Vector Machine. (Prihono and Sari, 2019).

Penelitian *Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending* memperkenalkan model topik untuk mengekstraksi fitur berharga dari teks deskriptif tentang pinjaman dan bangun empat model prediksi standar untuk diperagakan kinerja fitur-fitur ini untuk prediksi standar. (Jiang et al., 2018).

Dengan menggunakan database relasional, bahasa query terstruktur (SQL) dan analisis teks, kami membuat beberapa temuan: (i) mayoritas P2PL di Google Play tidak sah; (ii) rata-rata, P2PL resmi menerima peringkat ulasan yang lebih baik; (iii) ada banyak ulasan negatif terkait dengan metode penagihan hutang yang tidak etis dan pengenaan bunga yang berlebihan; dan (iv) ada empat P2PL yang memerlukan pengawasan khusus dari Otoritas Jasa Keuangan (OJK). Selain itu, sesuai dengan temuan, OJK tidak boleh secara pasif menunggu laporan resmi diajukan oleh publik terkait pelanggaran bisnis P2PL (Pranata and Farandy, 2019).

Penelitian dalam judul *A new model for software defect* Penelitian yang berjudul Analisis Sentimen Terhadap Review Fintech Dengan Metode Naive Bayes Classifier Dan K- Nearest Neighbor, dengan banyaknya review yang ditampilkan pada kolom komentar yang telah disediakan oleh Google Play Store di Aplikasi Dana dibutuhkan analisis untuk mengklasifikasi ulasan yang diberikan termasuk positif atau negatif. (Surohman et al., 2020).

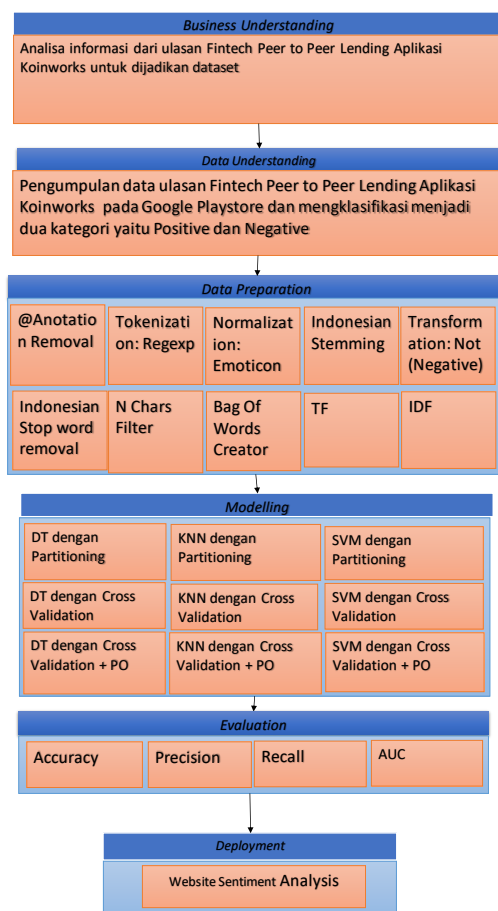
Penelitian ini akan menganalisis mengenai permasalahan yang berkaitan dengan beberapa ulasan tentang *Fintech* P2PL aplikasi Koinworks pada ulasan di *Google Play Store* serta menentukan hasil akurasi Analisis sentimen yang dihasilkan algoritma *Decision Tree*, *K-Nearest Neighbor* dan *Support Vector Machine*. Adapun manfaat dari penelitian ini adalah untuk membantu manajemen aplikasi Koinworks mengenai opini positif atau negatif dari pengguna aplikasi serta dapat memberikan bukti

secara empiris untuk teori yang berkaitan sehingga dapat dijadikan sumbangan pemikiran untuk pengembangan teori berikutnya.

2. METODE PENELITIAN

Dalam mendesain metode penelitian eksperimen menggunakan metode penelitian standar yang digunakan pada data mining yaitu *Cross-industry Standard Process for Data Mining* (CRISP-DM) yaitu terdiri dari 6 tahap dengan langkah-langkahnya adalah *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, dan *Evaluation*.

Berikut ini pada Gambar 2 di bawah ini menunjukkan model yang diusulkan dalam penelitian dengan menggunakan metode CRISP-DM:



Gambar 1. Model Penelitian Menggunakan CRISP-DM

2.1. Pemahaman Bisnis (*Business Understanding*)

Pada tahap ini dilakukan pemahaman terhadap objek penelitian yang dilakukan dengan menemukan informasi melalui ulasan pengguna *Google Play Store* tentang *Financial Technology Peer To Peer Lending* aplikasi koinworks yang mengekspresikan berbagai macam pendapat baik negatif maupun positif pada kolom ulasan yang disediakan.

Business Understanding juga dilakukan untuk pemahaman dalam mencari metode dengan pendekatan model analisis sentimen terbaik agar

dapat membantu untuk menentukan model yang tepat untuk digunakan berdasarkan perbandingan hasil algoritma. Adapun algoritma yang digunakan yaitu DT, KNN, dan SVM.

2.2. Pemahaman Data (*Data Understanding*)

Pada tahap *Data Understanding* bertujuan untuk mengumpulkan, mengidentifikasi, dan memahami data yang dimiliki. Data tersebut juga harus dapat diverifikasi kebenarannya. Data yang akan digunakan dalam penelitian ini adalah data dari ulasan aplikasi *fintech peer to peer lending* pada aplikasi Koinworks yang terdiri dari kategori positif dan negatif. Data yang digunakan dalam penelitian ini adalah data ulasan terhadap aplikasi Koinworks di *Google Play Store*.

2.3. Persiapan Data (*Data Preparation*)

Tahap *Data Preparation* diperoleh dari *crawling* data menggunakan aplikasi Webharvy dari ulasan aplikasi *fintech peer to peer lending* pada aplikasi Koinworks yang terdiri dari kategori positif dan negatif. Kemudian dilakukan *cleansing* data untuk mengurangi seperti data yang redundan kemudian *tokenizing*, normalisasi kata, *stopword removal*, dan *stemming*.

2.4. Pemodelan (*Modelling*)

Dalam tahap *Modelling* ini akan dilakukan teknik pengklasifikasian data yang paling akurat. Untuk membandingkan atau mengkomparasi, pada penelitian ini akan digunakan algoritma *Decision Tree*, *K Nearest Neighbor* dan *Support Vector Machine*.

2.5. Evaluasi (*Evaluation*)

Model yang terbentuk akan diuji menggunakan *confusion matrix* yang akan mengetahui tingkat akurasi. *Confusion Matrix* akan menggambarkan hasil akurasi mulai dari prediksi positif yang benar, prediksi positif yang salah, prediksi negatif yang benar dan prediksi negatif yang salah. Akurasi akan dihitung dari seluruh hasil prediksi yang benar (baik prediksi positif dan negatif) dibandingkan dengan seluruh data testing. Semakin tinggi nilai akurasi, semakin baik pula model yang dihasilkan. Pengujian juga diukur dengan menggunakan *ROC Curve*. *ROC Curve* akan menggambarkan kelas positif dalam bentuk kurva. Pengujian dilakukan dengan menghitung nilai *Area Under Curve* (AUC), semakin tinggi nilai AUC dan *ROC Curve*, maka semakin baik pula model klasifikasi yang terbentuk.

2.6. Penerapan (*Deployment*)

Tahap ini adalah tahap terakhir dari CRISP-DM, yaitu hasil dari seluruh tahapan yang sebelumnya digunakan secara nyata. Maknanya adalah melakukan sesuatu berdasarkan pengetahuan yang didapatkan dari kegiatan mining terhadap data. Penerapan dalam penelitian ini akan dikembangkan dengan PHP dan MySQL.

3. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah berasal dari kolom ulasan aplikasi Koinworks yang ada di website *Google Play Store*. *Google Play Store* adalah aplikasi resmi milik Google untuk perangkat yang menggunakan sistem baik operasi Android maupun Web. Bagi pengguna bisa mendapatkan aplikasi Koinworks dapat dengan cara mengunduh di *Google Play Store*. Pada *Google Play Store* juga terdapat kolom ulasan yang berisi ulasan para pengguna. Banyaknya dan beragam ulasan tersebut dapat dijadikan sebagai bahan penelitian di bidang *data mining*.

Cara untuk mendapatkan ulasan tersebut dilakukan proses yang di namakan *Web Scraping*. *Web Scraping* adalah teknik untuk pengambilan sebuah dokumen dari internet umumnya berupa halaman web yang dapat diambil datanya sesuai kebutuhan. Manfaat *web scraping* adalah memperoleh data melalui pengambilan dan ekstraksi data dengan ukuran data yang bervariasi. Data diambil dari bulan Januari 2017 sampai dengan Juni 2020.

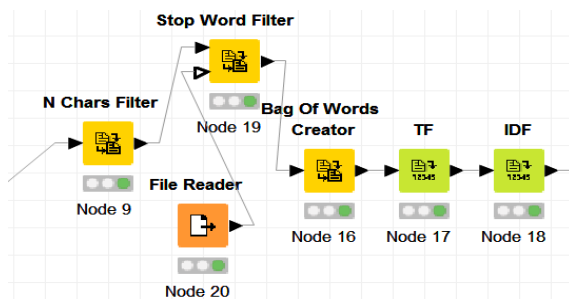
Hasil dari *cleansing* data yang digunakan dalam menganalisis sentimen aplikasi Koinworks dengan total 1560 .

Tabel 1. Tabel Hasil Proses Pelabelan

Label	Jumlah
Positive	798
Negative	762

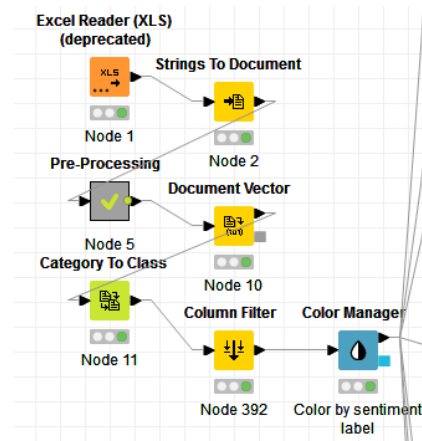
Tahap *data preparation* merupakan tahap dengan proses penyiapan data yang bertujuan untuk mendapatkan data yang bersih dan siap untuk digunakan dalam penelitian. Dalam *text mining* tahapan awal yang akan dilakukan adalah tahap *text preprocessing*, pada tahap ini peneliti menggunakan GATA framework dan tools KNIME. Data preparation yang digunakan pada web GATA framework adalah *Indonesian Stop Word Removal*, *Indonesian Stemming*, *Remove URL*, *@Anotation Removal*, *Indonesian Slank*, *Normalization Emoticon* dan *Transformation (Not Negative)*.

Berikut model *pre-processing* di tools Knime yang menggunakan simpul *N Chars Filter*, *Bag Of Words Creator* dan *TF IDF*.



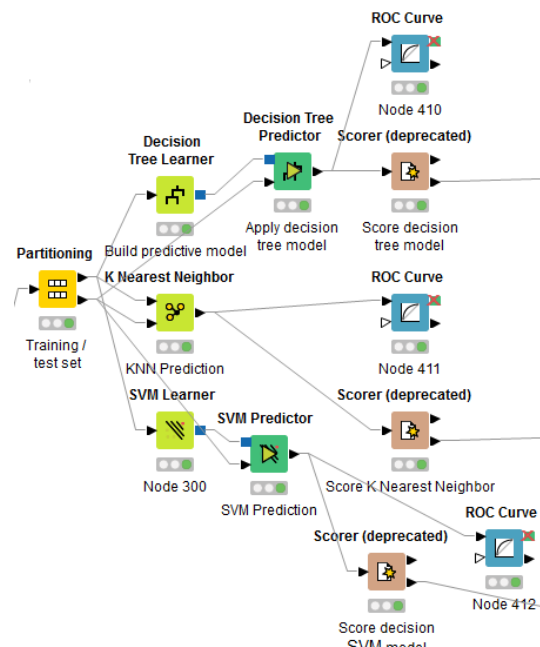
Gambar 2. Desain Model *Preprocessing* Data Menggunakan Knime

Gambar 2. menunjukkan desain model *preprocessing* dengan KNIME, data dengan tahap yang dilakukan adalah *case converter*, *RegEx Filter* (*@annotation removal*, *hashtag removal*, *remove exclude a-z*), *N Chars Filter*, Tokenisasi (*Bag of Words Creator*), TF, dan IDF.



Gambar 3. Desain Modeling Algoritma DT, KNN, dan SVM

Gambar 3. menunjukkan *desain* model algoritma DT, KNN dan SVM untuk ketiga kategori yaitu Partitioning, Cross Validation dan Parameter Optimazion, *desain* model tersebut diawali dengan operator *Excel Reader (XLS)*, *Strings To Document*, *Preprocessing*, *Document Vector*, *Category To Class*, *Color Manager*, *Partitioning*, dan *node analytics*.



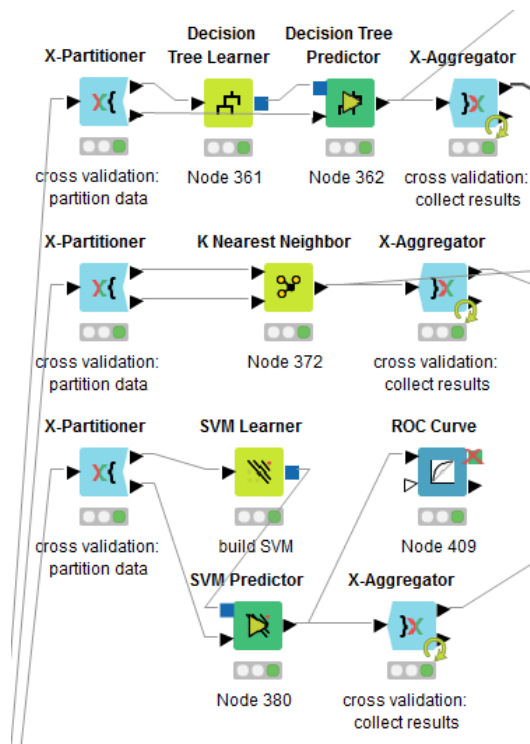
Gambar 4. Algoritma DT, KNN dan SVM dengan Partitioning

Partitioning merupakan Tabel input dibagi menjadi dua partisi (yaitu berdasarkan baris), misal melatih dan menguji data. Kedua partisi tersedia di dua port output. Cross Validation Cross Validation adalah teknis validasi yang membagi data ke dalam K bagian dan kemudian masing-masing bagian akan

dilakukan proses klasifikasi. Model Model Parameter Optimazion digunakan untuk mengoptimalkan dan mengkonfigurasi parameter terbaik.

Modelling merupakan fase pemilihan teknik mining dengan menentukan algoritma yang digunakan. Dalam tahap *Modelling* ini dilakukan teknik pengklasifikasian data dengan membandingkan tiga algoritma yaitu *Decision Tree*, *K Nearest Neighbor* dan *Support Vector Machine* menggunakan *Partitioning*, *Cross Validation*, serta *Cross Validation + Parameter Optimize*.

Pada Gambar 4. merupakan modelling algoritma *Decision Tree*, *K Nearest Neighbor* dan *Support Vector Machine* dengan *Partitioning*.



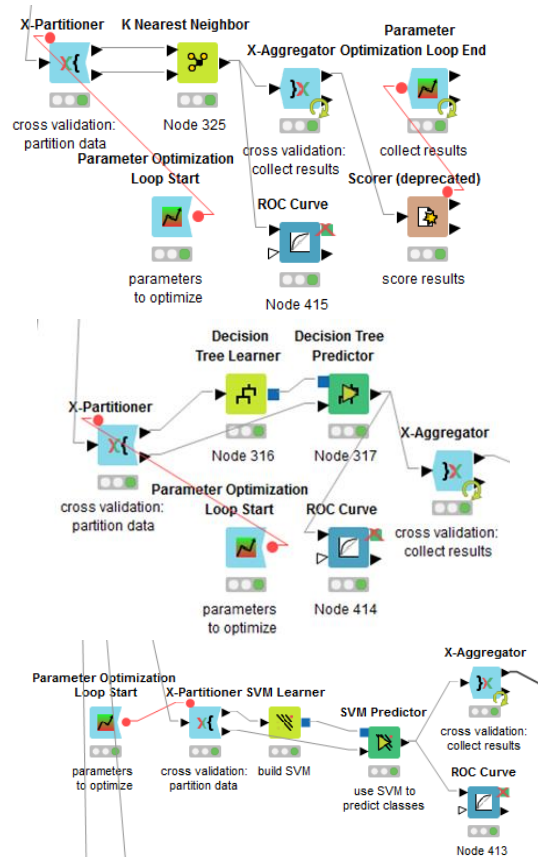
Gambar 5.. Algoritma DT, KNN dan SVM dengan Cross Validation

Pada Gambar 5. merupakan modelling algoritma *Decision Tree*, *K Nearest Neighbor* dan *Support Vector Machine* dengan *Cross Validation*.

Pada Gambar 6. modelling algoritma *Decision Tree*, *K Nearest Neighbor* dan *Support Vector Machine* dengan *Cross Validation + Parameter Optimize*.

Adapun perbandingan hasil *accuracy*, *precision*, *recall* dan AUC Algoritma DT, KNN dan SVM dengan *Partitioning*, *Cross Validation*, dan *Parameter Optimize* telah digunakan sebagai berikut :

Berdasarkan Tabel 2. Hasil perhitungan metode DT dengan *partitioning* mendapatkan nilai akurasi 85,47% perhitungan metode KNN dengan *partitioning* dengan $k = 3$ menghasilkan nilai akurasi 82,91% sedangkan hasil perhitungan SVM dengan *partitioning* mendapatkan nilai akurasi 88,03%.



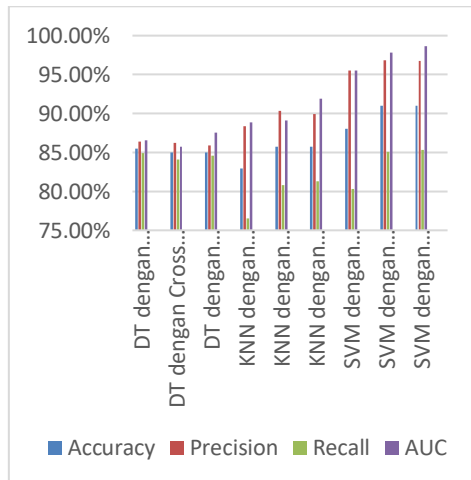
Gambar 6.. Algoritma DT, KNN dan SVM dengan Cross Validation + Parameter Optimize

Hasil perhitungan metode DT dengan *cross validation* mendapatkan nilai akurasi 85,00% perhitungan metode KNN dengan *cross validation* dengan $k = 3$ menghasilkan nilai akurasi 85,77% sedangkan hasil perhitungan SVM dengan *cross validation* mendapatkan nilai akurasi 90,96%.

Tabel 2. Perbandingan Accuracy, Precision, Recall dan AUC

Algoritma	P	Accuracy	Precision	Recall	AUC
DT + Partitioning	P	85,47%	86,38%	84,94%	0,865
DT + Cross Validation	N	85,00%	86,25%	84,09%	0,858
DT + Parameter Optimize	P	85,00%	85,88%	84,59%	0,876
KNN + Partitioning	P	82,91%	88,41%	76,57%	0,888
KNN + Cross Validation	N	82,91%	88,41%	76,57%	0,888
KNN + Parameter Optimize	P	85,77%	90,34%	80,83%	0,891
SVM + Partitioning	P	88,03%	95,52%	80,33%	0,956
SVM + Cross Validation	N	90,96%	96,86%	85,09%	0,979
SVM + Parameter Optimize	P	91,03%	96,73%	85,34%	0,986

Hasil perhitungan metode DT dengan *parameter optimize* mendapatkan nilai akurasi 85,00%, perhitungan metode KNN *parameter optimize* dengan $k = 3$ menghasilkan nilai akurasi 85,77% sedangkan hasil perhitungan SVM dengan *parameter optimize* mendapatkan nilai akurasi 91,03% .

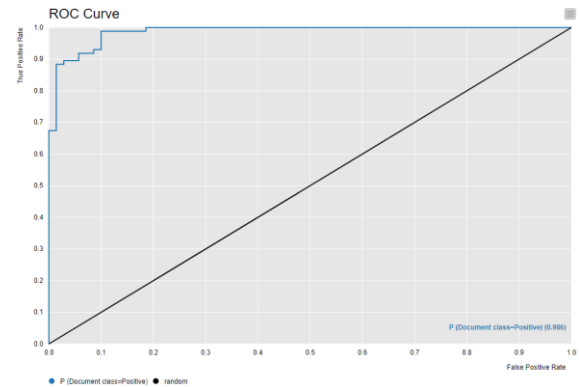


Gambar 7. Grafik Perbandingan Accuracy, Precision, Recall dan AUC

Berdasarkan Gambar 7. dapat disimpulkan bahwa analisis sentimen dengan menggunakan metode SVM *Cross Validation + Parameter Optimize* menghasilkan nilai *accuracy*, *precision* terbaik dibandingkan dengan DT Partitioning, KNN Partitioning, SVM Partitioning, DT Cross Validation, KNN Cross Validation, SVM Cross Validation, DT Cross Validation + *Parameter Optimize* dan KNN Cross Validation + *Parameter Optimize*. Penggunaan metode SVM dengan *Cross Validation + Parameter Optimize* menghasilkan AUC terbaik yaitu 0,986 sehingga termasuk **excellent classification**. Sehingga dapat disimpulkan bahwa akurasi C mendapatkan akurasi yang sama dan lebih tinggi dibandingkan dengan DT dan KNN yang menggunakan *partitioning*, *cross validation*, dan *parameter optimize*. Metode SVM dengan *Cross Validation + Parameter Optimize* mendapatkan hasil yang terbaik karena dengan penggunaan parameter optimize dapat di sesuaikan parameter yang digunakan, pada penelitian ini menggunakan parameter maksimal number of records dengan start value=1 dan stop value=7 dan step size=3.

Berikut ini merupakan Kurva ROC dan *Confusion Matrix* dari algoritma SVM dengan *Cross Validation Parameter Optimize* dapat dilihat pada Gambar 8.

Berdasarkan Gambar 8. Kurva ROC SVM dengan *Cross Validation* dan *Parameter Optimize* dengan nilai AUC (Area Under Curve) yang dihasilkan dari gambar 4.36 di atas sebesar 0,986 dimana diagnosa hasilnya adalah **excellent classification**.

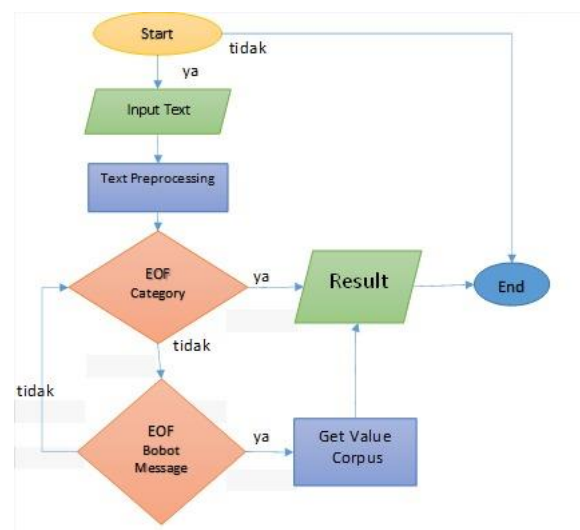


Gambar 8. Kurva ROC SVM dengan *Cross Validation* dan *Parameter Optimize*

Model klasifikasi teks memudahkan untuk mengetahui klasifikasi sentimen positif dan negatif. Berdasarkan *dataset* yang diolah menggunakan *KNIME*, data ulasan terpisah menjadi kata-kata yang memiliki bobot pada setiap kata-katanya. Kata-kata tersebut digunakan untuk melihat kata-kata yang berhubungan dengan sentimen yang sering muncul dan memiliki bobot tertinggi dan dapat digunakan untuk mengetahui *opini public* terhadap *fintech peer to peer lending* pada aplikasi Koinworks.

Setelah didapatkan *korpus data* dari hasil evaluasi maka *korpus data* tersebut dapat dimasukkan ke dalam *database* untuk proses *deployment* aplikasi analisis sentimen *fintech P2PL* aplikasi Koinworks.

Berikut di bawah merupakan *flowchart* dari aplikasi analisis sentimen *fintech P2PL* aplikasi Koinworks yang dibuat pada Gambar 9. Berdasarkan *flowchart* pada Gambar 9. dapat dilihat bahwa dalam aplikasi Analisis Sentimen *Fintech P2PL* Aplikasi Koinworks cara dalam melakukan input analisis teks yaitu dengan diinput secara manual yang kemudian akan dilakukan proses *text preprocessing*.



Gambar 9. Flowchart Aplikasi Analisis Sentimen Fintech P2PL Aplikasi Koinworks



Gambar 10. Halaman Data dan Grafik Hasil Sentimen

Pada gambar 10. merupakan menu utama sekaligus menu *dashboard* yang menampilkan *pie chart* jumlah perbandingan kategori *positive* dan *negative* yang sudah dilakukan analisis di bagian kiri halaman. Pada bagian kanan halaman terdapat tabel yang menunjukkan hasil analisis sentimen fintech P2PL aplikasi Koinworks, analisis tersebut berasal dari input secara manual, hasil prediksi kategori beserta dengan bobot nilai per masing-masing kategori.

4. KESIMPULAN

Berdasarkan hasil penelitian dapat ditarik kesimpulan bahwa bahwa penerapan *data mining* dapat membantu mengkategorikan sentimen ulasan *fintech* P2PL pada aplikasi Koinworks. Algoritma SVM dengan penambahan *Cross Validation* dan *Parameter Optimize* merupakan metode yang terbaik dalam penelitian ini dibandingkan dengan DT, KNN baik *partitioning*, *Cross Validation* maupun *Parameter Optimize* untuk melakukan klasifikasi sentimen analisis teks berbahasa Indonesia dalam analisis sentimen terhadap *fintech peer to peer lending* pada aplikasi Koinworks. Algoritma SVM dengan *Cross Validation + Parameter Optimization* menghasilkan *Accuracy* 91,03% *precision* tertinggi yaitu dengan 96,73% , *recall* 85,34% dan *AUC* tertinggi yaitu 0,986 yang termasuk dalam *excellent classification*. Usulan penggunaan *Cross Validation + Parameter Optimization* pada algoritma klasifikasi terbukti berpengaruh meningkatkan *accuracy* algoritma SVM namun tidak secara signifikan. Sehingga dapat disimpulkan bahwa penerapan optimasi yang terbaik pada model ini adalah algoritma SVM dengan *cross validation + parameter optimize* dapat memberikan solusi terhadap permasalahan klasifikasi sentimen *fintech* P2PL pada aplikasi Koinworks.

DAFTAR PUSTAKA

- AAPUTRA, S.A., DIDI ROSIYADI, WINDU GATA AND SYEPRY MAULANA HUSAIN, 2019. Sentiment Analysis Analisa Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 3(3), pp.377–382.
- AJI, S., WARJIYONO, W., PRATMANTO, D., ARDIANSYAH, A., WIDODO, A., FAQIH, H., SULEMAN, S. AND FANDHILAH, F., 2019. Review Sentiment Analysis of World Class Hotel Using Naive Bayes Classifier And Particle Swarm Optimization Method. [online] (January 2018). Available at: <https://www.researchgate.net/publication/330792673_Review_Sentiment_Analysis_of_World_Class_Hotel_Using_Naive_Bayes_Classifier_And_Particle_Swarm_Optimization_Method/link/5c885efb299bf14e7e782e61/download>.
- BALAKRISHNAN, V., SELVANAYAGAM, P.K. AND YIN, L.P., 2020. Sentiment and Emotion Analyses for Malaysian Mobile Digital Payment Applications. [online] pp.67–71. Available at: <<https://dl.acm.org/doi/pdf/10.1145/3388142.3388144>>.
- CHANGE.ORG, 2018. *Penagihan Pinjaman Fintech Sangat Meresahkan*. [online] Available at: <<https://www.change.org/p/ojk-penagihan-pinjaman-fintech-sangat-meresahkan>>.
- IGNATOW, G. AND MIHALCEA, R., 2015. *An introduction to text mining. Research Design, Data Collection, and Analysis*.
- JIANG, C., WANG, Z., WANG, R. AND DING, Y., 2018. Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. *Annals of Operations Research*, 266(1–2), pp.511–529.
- KURNIAWAN, T.A., WARDANI, D.K. AND WIDHAYATI, L., 2019. Pengaruh Keberterimaan Layanan Peer To Peer Lending Kepada Umkm Sebagai Pengguna Dengan Menggunakan Metode Technology Acceptance Model (Tam). *Jurnal Sosial Ekonomi Dan Humaniora*, [online] 5(2), pp.151–160. Available at: <<http://www.jseh.unram.ac.id/index.php/jseh/article/view/59>>.
- MAHENDRAJAYA, R., BUNTORO, G.A. AND SETYAWAN, M.B., 2019. Analisa Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based Dan Support Vector Machine. *Komputek*, [online] 3(2), p.52. Available at: <<http://studentjournal.umpo.ac.id/index.php/komputek/article/viewFile/270/246>>.
- NARAIN, A., 2017. *Two Faces of Change*. Finance & Development.
- NUGRAHA, A.P., ROLANDO, PUSPASARI, M.A. AND SYAIFULLAH, D.H., 2019. Usability Evaluation for User Interface Redesign of Financial Technology Application. *IOP Conference Series: Materials Science and Engineering*, [online] 505(1). Available at: <<https://iopscience.iop.org/article/10.1088/1757-899X/505/1/012101/pdf>>.
- OJK, 2020. *Perkembangan Fintech Lending*. [online] Available at:

- <[https://www.ojk.go.id/id/kanal/iknb/data-dan-statistik/fintech/Documents/Perkembangan Fintech Lending Periode Maret 2020.pdf](https://www.ojk.go.id/id/kanal/iknb/data-dan-statistik/fintech/Documents/PerkembanganFintechLendingPeriodeMaret2020.pdf)>.
- PRANATA, N. AND FARANDY, A.R., 2019. Big Data-Based Peer-to-Peer Lending Fintech: Surveillance System through Utilization of Google Play Review. *SSRN Electronic Journal*, (943).
- PRIHONO, O.F. AND SARI, P.K., 2019. Comparison Analysis Of Social Influence Marketing For Mobile Payment Using Support Vector Machine. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, [online] 4(4), pp.367–374. Available at: <<https://kinetik.umm.ac.id/index.php/kinetik/article/view/921/pdf>>.
- PRIYADHARSINI.C AND THANAMANI, D.A.S., 2014. An Overview of Knowledge Discovery Databaseand Data mining Techniques. *International Journal of Innovative Research in Computer and Communication Engineering*, [online] 2(1), pp.1571–1578. Available at: <<http://www.rroij.com/open-access/an-overview-of-knowledge-discovery-databaseand-data-mining-techniques.php?aid=48833>>.
- RAMADHANI, N., 2019. *10 Platform P2P Lending Terbaik Versi KPMG di Fintech Edge*. [online] www.akseleran.co.id. Available at: <<https://www.akseleran.co.id/blog/fintech-edge/>>.
- SUDIRA, H., DIAR, A.L. AND RULDEVIYANI, Y., 2019. Instagram Sentiment Analysis with Naive Bayes and KNN: Exploring Customer Satisfaction of Digital Payment Services in Indonesia. *2019 International Workshop on Big Data and Information Security, IWBIS 2019*, [online] pp.21–26. Available at: <<https://ieeexplore.ieee.org/abstract/document/8935700/>>.
- SUROHMAN, AJI, S., ROUSYATI AND WATI, F.F., 2020. Analisa Sentimen Terhadap Review Fintech Dengan Metode Naive Bayes. *Evolusi: Jurnal Sains dan Manajemen*, [online] 8(1), pp.93–105. Available at: <<https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/7535/4065>>.
- WISNU, H., AFIF, M. AND RULDEVYANI, Y., 2020. Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naive Bayes. *Journal of Physics: Conference Series*, [online] 1444(1). Available at: <<https://iopscience.iop.org/article/10.1088/1742-6596/1444/1/012034/PDF>>.
- WULANDARI, P.P., 2017. Analisa Faktor-Faktor Yang Menentukan Keputusan Pemberian Kredit Untuk Usaha Mikro, Kecil, Dan Menengah (Umk) Pada Lembaga Pembiayaan Peer To Peer Lending. *Journal of Chemical Information and Modeling*, [online] 53(9), pp.1689–1699. Available at: <<file:///C:/Users/User/Downloads/fvm939e.pdf>>.

Halaman ini sengaja dikosongkan