

BAB I PENDAHULUAN

1.1. Latar Belakang

Saham dapat diartikan sebagai tanda bukti kepemilikan modal/dana pada suatu perusahaan. Sebuah kertas yang tercantum dengan jelas nilai nominal, nama perusahaan dan di ikuti dengan hak dan kewajiban yang dijelaskan kepada setiap pemegangnya dan juga merupakan persediaan yang siap untuk dijual (Umaidah, 2018). Pasar modal adalah tempat dimana berbagai pihak khususnya perusahaan menjual saham (stock) dan obligasi (bond) dengan tujuan dari hasil penjualan tersebut yang nantinya akan dipergunakan sebagai tambahan dana atau untuk memperkuat modal perusahaan (Riyanto, 2017). Dalam setiap transaksi pasar modal, seorang pialang dihadapkan dengan pilihan menjual atau membeli saham. Meskipun dapat di prediksi dengan analisis teknikal maupun analisis fundamental, tetapi kesalahan sering terjadi dalam proses pembacaan dan pengkolerasian nilai saham (Roy, 2016). Dalam melakukan peramalan harga saham sulit untuk mengetahui titik waktu terbaik untuk membeli atau menjual saham, karena banyak variabel yang dapat mempengaruhi pasar saham (Su and Cheng, 2016). Untuk memperoleh hasil prediksi yang akurat terhadap pengembalian stok pasar saham merupakan tugas yang sulit karena tidak ratanya rangkaian waktu keuangan, karakteristik data *nonlinier*, besarnya ketidakpastian dan kompleksitas (Qiu *et. al.*, 2016). Analisis fundamental digunakan untuk menentukan jenis saham yang akan investor beli (Hakmi, Rahayu and Topowijono, 2017). Sedangkan analisis teknikal digunakan untuk memprediksi pergerakan saham berdasarkan pola harga saham dengan menggunakan indikator pada jenis saham tertentu (Abidin, Suhadak and Hidayat, 2016). Variabel prediksi yang biasa digunakan dalam literatur untuk model SVM (*Support Vector Machine*) adalah indikator-indikator analisis teknikal, indikator berasal dari fungsi dan rumus tertentu untuk harga saham masa lalu (Henrique *et. al.*, 2018a). Adapun jenis indikator pada analisis teknikal yaitu XMA (*Exponential Moving Average*). STI (*Stochastic Oscillator*).

RSI (*Relative Strength Index*). PROC (*Rate of change and momentum*). ADO/CLV (*Accumulation/Distribution*), apabila salah dalam membaca indikator, maka kerugian besar akan dialami seorang pialang (Henrique, Sobreiro and Kimura, 2018b). Penambangan *outlier* telah diterapkan di pasar saham, data yang dianggap memiliki sifat yang berbeda dibandingkan dengan kebanyakan data lainnya. Data *outlier* tetap digunakan karena mengandung beberapa informasi penting (Li *et. al.*, 2016). Untuk memprediksi harga saham dengan akurasi tinggi menjadi sangat sulit, walaupun sudah dilakukan pengurangan variabel yang tidak penting tetapi masih tetap adanya *outlier* (Jeon *et. al.*, 2017). Teknik pengelompokan data sangat banyak digunakan pada penambangan data, teknik pengelompokan *K-Means* tidak dapat menghasilkan *cluster* berkualitas baik dengan data dimensi tinggi, jika diamati dari standar deviasi dan struktural kesalahan SOM (*Self Organizing Map*) lebih kecil (Nair *et. al.*, 2016). Namun masih sulit mengetahui nilai kelas paling optimal pada tahapan *cluster* data (Bansal *et. al.*, 2017). Untuk menyelesaikan masalah tersebut diperlukan sebuah model yang dapat menyelesaikan prediksi trend harga saham berdasarkan indikator-indikator analisis teknikal.

Beberapa peneliti melakukan perbandingan kerangka kerja SVM (*Support Vector Machine*) – KNN (*Knearest Neighbor*) dengan FWSVM (*Feature Weighted Support Vector Machine*) dan FWKNN (*Feature Weighted Knearest Neighbor*) untuk dua indeks pasar saham dan dipengaruhi oleh jumlah variabel fitur, dari hasil yang diperoleh menunjukkan bahwa FWSVM-FWKNN lebih baik (Chen dan Hao, 2017). Dilakukan juga penelitian dengan menggunakan KNN yang telah di modifikasi, yaitu dengan EEMD (*Ensemble Empirical Mode Decomposition*) – MKNN (*Multidimensional k-Nearest Neighbor*). Hasil yang di peroleh menunjukkan bahwa model EEMD-MKNN memiliki ketepatan hasil perkiraan yang lebih tinggi dibandingkan dengan KNN, yaitu MAPE EEMD-MKNN 0,4997 dan MAPE KNN 0,5955 (Zhang *et. al.*, 2017). Sebelumnya juga sudah dilakukan penelitian prediksi harga saham oleh (Sinaga, *et. al.*, 2019) dengan menggunakan teknik SV-kNNC (*Support Vector - k Nearest Neighbor*

Clustering) dan SOM (*Self Organizing Map*) untuk memprediksi harga saham naik dan turun dengan adanya perbandingan pada tahapan *clustering*. Hasil penelitian tersebut memperlihatkan *SV-kNNC* dan *SOM* adalah model yang efektif untuk memprediksi arah pergerakan harga saham naik atau turun dan kinerjanya lebih baik dari pada *SV-kNNC* dan *K-Means*. Namun pada penelitian ini untuk penentuan nilai kelas paling optimal pada tahapan *cluster* masih dilakukan dengan cara manual yaitu dengan menguji nilai dari kelas data satu per satu. Pada tahapan klasifikasi pada penelitian ini dilakukan dengan teknik KNN (*K-Nearest Neighbors*) untuk memperoleh hasil prediksi harga saham, namun akurasi yang diperoleh dari teknik klasifikasi ini kurang maksimal.

Untuk mengatasi masalah dari model *SV-kNNC* dan *SOM* maka perlu dilakukan optimasi pada tahapan *clustering* untuk memperoleh nilai kelas terbaik pada saat pengelompokan data saham secara otomatis. (Yuan dan Yang, 2019) mengusulkan metode *Silhouette Coefficient*, teknik ini menggunakan kuadrat jarak antara titik sampel di setiap *cluster* dan pusat massa *cluster* untuk memberikan serangkaian nilai K untuk meningkatkan kinerja *SOM* pada tahapan *cluster* data. Dari model *SV-kNNC* dan *SOM* juga perlu dilakukan optimasi pada tahapan klasifikasi data untuk meningkatkan nilai akurasi prediksi harga saham. Untuk mengatasi nilai akurasi prediksi saham (Mehta *et. al.*, 2018) mengusulkan metode klasifikasi LMKNN (*Local Mean-Based K Nearest Neighbor*) metode yang disebut K-tetangga terdekat berbasis rata-rata telah diusulkan, metode ini adalah klasifikasi nonparametrik sederhana dan efektif. LMKNN ini telah terbukti meningkatkan kinerja klasifikasi dan juga mengurangi efek *outlier* yang ada, terutama dalam ukuran data yang kecil.

Dari permasalahan dan kajian literatur yang dilakukan maka diusulkan sebuah model menggunakan *SV-kNNC-SilhouetteCoefficient-LMKNN* untuk mengatasi penentuan nilai kelas pada tahapan *cluster* data yang masih manual, dan juga untuk meningkatkan hasil akurasi prediksi harga saham pada tahapan klasifikasi.

Berdasarkan latar belakang tersebut, maka akan dilakukan penelitian dengan judul “**Optimisasi SV-kNNC dengan menggunakan *SilhouetteCoefficient* dan LMKNN untuk Prediksi Harga Saham**”.

1.2. Masalah Penelitian

Berdasarkan uraian pada latar belakang di atas, maka masalah dalam penelitian ini dapat diuraikan menjadi dua bagian yaitu identifikasi masalah dan rumusan masalah.

1.2.1 Identifikasi Masalah

Adapun identifikasi masalah pada latar belakang di atas adalah sebagai berikut:

1. Sulitnya mengetahui kelas terbaik pada tahapan pengelompokan data untuk prediksi harga saham.
2. Kurang optimalnya nilai akurasi pada tahapan klasifikasi pada saat melakukan prediksi harga saham.

1.2.2 Rumusan Masalah

Adapun rumusan masalah dalam penelitian tesis ini adalah bagaimana menghasilkan model untuk penentuan nilai kelas terbaik pada tahapan pengelompokan data yang masih manual dan bagaimana menghasilkan model untuk meningkatkan akurasi prediksi harga saham pada tahapan klasifikasi.

1.3. Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah menghasilkan model yang dapat memberikan nilai K di awal pada tahapan pengelompokan data dan meningkatkan akurasi prediksi harga saham pada tahapan klasifikasi.

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Model ini dapat digunakan oleh pialang untuk menentukan strategi trading, sehingga dapat menjadi penunjang keputusan seorang trader/pialang dalam mengambil keputusan membeli atau menjual saham.

2. Model ini dapat dapat digunakan sebagai referensi untuk penelitian sejenis lainnya.

1.4. Batasan Masalah

Adapun batasan dari penelitian ini adalah sebagai berikut:

1. Dataset yang akan diprediksi terdaftar dibursa efek Indonesia yang memiliki historical data sepanjang 5 tahun dari <https://finance.yahoo.com/> diawali dari tahun 2015 hingga 2020.
2. Jenis indikator yang digunakan adalah indikator pada analisis teknikal.
3. Indikator-indikator saham yang digunakan sebagai pembanding nilai saham yaitu XMA (*Exponential Moving Average*)10. XMA (*Exponential Moving Average*)20. XMA (*Exponential Moving Average*)30. STI (*Stochastic Oscillator*)14. STI (*Stochastic Oscillator*)19. STI(*Stochastic Oscilator*)30. RSI (*Relative Strength Index*)14. RSI(*Relative Strength Index*)19. PROC (*Rate of change and momentum*). ADO/CLV (*Accumulation/Distribution*). Setiap nomor pendamping indikator menjelaskan berapa lama hari yang digunakan datanya untuk proses perhitungan.
4. Perusahaan yang digunakan dalam penelitian ini yaitu perusahaan *Blue chip*. perusahaan *Blue chip* merupakan perusahaan yang memiliki pendapatan stabil dan memiliki aset yang besar. Ada lima perusahaan yang peneliti buat yaitu: Bank Rakyat Indonesia (BBRI). Astra Internasional (ASII JK). Perusahaan gas Negara (PGAS). Telkom Indonesia (TLKM) dan Bank Cental Indonesia (BBCA).
5. Data yang digunakan perusahaan ASII JK sebanyak 1376 data. data yang digunakan perusahaan BBRI JK sebanyak 1376 data. data yang digunakan perusahaan PGAS JK sebanyak 1376. data yang digunakan perusahaan TLKM JK sebanyak 1376 data. data yang digunakan perusahaan BBCA JK sebanyak 1376 data.
6. Hasil dari model yang dihasilkan merupakan trend naik dan turun dari prediksi nilai harga saham.

7. Menggunakan *Dynamic Link Library* (DLL) untuk hasil metode *Support Vector Machine* (SVM).

1.5. Metodologi Penelitian

Metodologi yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Studi Literatur

Pada tahap ini dilakukan proses untuk memahami bagaimana konsep analisis teknikal untuk mengetahui cara kerja dari masing-masing indikator yang digunakan, memahami proses klasifikasi pada LMKNN dan proses penentuan kelas pada tahapan pengelompokan data pada *SilhouetteCoefficient* dan memahami proses prediksi harga saham pada metode yang sudah ada.

2. Analisis Masalah

Pada tahap ini dilakukan analisis berdasarkan hasil studi literatur untuk mengidentifikasi masalah yang harus diselesaikan, data yang dibutuhkan, dan menentukan metode yang diusulkan untuk menyelesaikan masalah.

3. Perancangan Model

Perancangan model dimulai dari mengumpulkan dataset saham, dataset saham akan dinormalisasi dalam rentang -1 sampai +1, setelah data di normalisasi, maka akan dilakukan perhitungan indikator data saham berdasarkan analisis teknikal. Selanjutnya di terapkan SVM untuk memperoleh *Support Vector* dari setiap data saham dan menentukan kelas data dari setiap *Support Vector* yang telah di peroleh. Sebelum dilakukan pengelompokan data akan di terapkan terlebih dahulu *SilhouetteCoefficient* untuk memperoleh K terbaik, maka K terbaik ini akan digunakan pada tahapan pengelompokan data dengan menggunakan SOM. Setelah dilakukan pengelompokan data, setiap data saham akan diberikan bobot. *Support Vector* yang telah diberikan bobot akan diklasifikasi menggunakan LMKNN untuk prediksi harga saham.

4. Pengujian

1. Membandingkan SOM + *SilhouetteCoefficient* dengan SOM untuk melihat pengaruh SOM pada tahapan pengelompokan data dalam mengatasi penentuan nilai K yang masih manual.
2. Untuk membandingkan LMKNN dengan KNN pada tahapan prediksi harga saham, untuk melihat pengaruh LMKNN pada tahapan klasifikasi data. Pengujian ini di lihat dengan menggunakan *Confution Matrix* untuk melihat nilai Akurasi.
 - a. Menarik kesimpulan dari hasil pengujian.
 - b. Menyusun laporan tesis.

1.6. Sistematikan Penulisan

Sistematika penulisan laporan penelitian ini terdiri dari 5 bab, dimana secara garis besar masing-masing bab membahas hal – hal berikut ini. Bab 1 Pendahuluan, berisi penjelasan umum, masalah dan solusi yang sudah ada dan akan dilakukan. Bab 2 berisi studi literatur dan tinjauan pustaka terkait masalah dan metode yang berhubungan dengan penelitian yang akan dilakukan. Bab 3 Metodologi Penelitian, berisi identifikasi masalah, langkah-langkah dari metode yang diusulkan, data yang digunakan, alat-alat penelitian dan metode analisis. Bab 4 Hasil dan Pengujian, berisi hasil yang diperoleh dari model yang dibangun dan pengujian yang dilakukan. Bab 5 Kesimpulan dan Saran, berisi kesimpulan yang diperoleh dari hasil dan pengujian penelitian yang dilakukan dan saran yang dapat dilakukan untuk hasil yang lebih baik pada penelitian selanjutnya.