

## 1. Pendahuluan

### Latar Belakang

Saham merupakan bukti penyertaan modal dalam suatu perusahaan atau perseroan terbatas. Saham juga merupakan instrument investasi yang populer dan memiliki risiko lebih tinggi dibandingkan dengan instrument investasi lainnya, tetapi investasi saham memberikan keuntungan paling tinggi dari dividen ataupun kenaikan harga saham. Risiko saham yang paling umum adalah *capital lost*, yaitu harga saham yang mengalami penurunan harga dibandingkan dengan harga pada saat pembelian. Kerugian investasi saham akan semakin besar jika harga saham mengalami penurunan yang sangat besar. Tidak ada kepastian mengenai posisi pergerakan harga indeks suatu saham. Oleh karena itu, untuk dapat membantu investor mengambil keputusan yang tepat dalam bertransaksi di pasar saham perlu memprediksi pergerakan harga saham. Untuk memprediksi pergerakan harga saham sangat sulit karena banyak faktor yang harus diperhatikan mulai dari peristiwa politik, ekonomi, sosial, budaya, dan banyak faktor lainnya. Salah satu faktor yang mempengaruhi harga saham adalah berita ekonomi. Berdasarkan hal ini, peneliti melakukan penelitian untuk menganalisis prediksi harga saham menggunakan sentiment ahli ekonomi yang diekspresikan dalam sebuah berita ekonomi tentang saham. Penelitian ini bertujuan untuk menganalisis dampak dari sentiment yang dikeluarkan oleh para ahli yang memiliki pengetahuan akan saham sehingga menghasilkan akurasi prediksi harga saham.

Berdasarkan penelitian sebelumnya [1] melakukan penelitian memprediksi harga saham menggunakan sentiment analisis tetapi data yang digunakan dalam penelitian berbeda, data yang digunakan dalam penelitian yaitu opini publik yang dituangkan di sosial media khususnya di facebook dan twitter. Oleh karena itu, Penelitian ini menggunakan klasifikasi *sentiment analysis* untuk memprediksi harga saham. *Sentiment analysis* adalah salah satu teknik *machine learning* dengan cara memproses komputasional dalam mengidentifikasi dan mengategorikan atau mengklasifikasikan opini-opini dalam bentuk potongan teks, khususnya untuk mengukur maksud pembuat potongan teks terhadap topik tertentu dengan hasil yang sentimen adalah positif atau negatif [2]. Berdasarkan hal tersebut, hasil dari sentiment analisis dapat membantu dan mendukung calon investor untuk membuat keputusan dalam bertransaksi di pasar saham.

Pada penelitian ini algoritma yang digunakan untuk melakukan pembobotan kata adalah *TF-IDF* dilihat dari performansi dari algoritma *TF-IDF* pada penelitian sebelumnya [1], Dimana algoritma *TF-IDF* mempunyai nilai performansi akurasi terbaik senilai 75.4 % dibandingkan dengan algoritma *bag-of-word* yang hanya mempunyai nilai akurasi 73.8% menggunakan metode klasifikasi *logistic regression*. Metode klasifikasi yang digunakan adalah *support vector machine*. Karena pada penelitian [3] disebutkan bahwa proses klasifikasi menggunakan *Support Vector Machine* mendapatkan evaluasi akurasi dengan tingkat akurasi hingga 76% dari seluruh proses training dan tingkat akurasi ini lebih baik dari algoritma *Naïve Bayes Classifier* yang hanya stabil sampai dengan 75%.

### Topik dan Batasannya

Topik dan batasan pada penelitian ini adalah melakukan *analysis sentiment* untuk pergerakan harga IHSG. Dataset yang digunakan berisi berita ekonomi terkait dengan IHSG berbahasa Indonesia. Pada penelitian ini klasifikasi dan pembobotan fitur juga terbatas pada performa terhadap *support vector machine*, *Term Frequency-Inverse Document Frequency*.

### Tujuan

Tujuan dari penelitian ini untuk menganalisis pengaruh sentimen berita dengan perubahan harga saham, serta untuk mengetahui performansi metode klasifikasi *support vector machine* terhadap pengaruh penelitian *sentiment analysis* berita dengan pergerakan harga saham.

### Organisasi Tulisan

**Studi terkait** berisi teori yang mendukung terkait dengan topik yang dikerjakan yaitu, *sentiment analysis*, *support vector machine* dan *TF-IDF*. **Sistem yang dibangun** menjelaskan rancangan atau sistem yang dibangun pada penelitian ini. **Evaluasi** berisi dua sub-bagian, yaitu Hasil Pengujian dan Analisis Hasil Pengujian. **Kesimpulan** memuat kesimpulan dan Saran, Semua poin kesimpulan diambil dari hasil pengujian dan analisis hasil pengujian.