

Nama : Angga Fathan Rofiqy
NIM : G1401211006
Judul : News impact on stock price return via sentiment analysis
Sumber : Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. 2014. News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, 69, 14-23. doi: 10.1016/j.knosys.2014.04.022

Abstract

Artikel berita keuangan diyakini memiliki dampak pada perubahan harga saham. Pekerjaan sebelumnya memodelkan potongan berita dalam ruang kata-kantong (bag-of-words), yang menganalisis hubungan laten antara pola statistik kata dan perubahan harga saham. Namun, sentimen berita, yang merupakan elemen penting dalam hubungan antara pola kata dan pergerakan harga, jarang dipertimbangkan. Dalam makalah ini, kami pertama-tama menerapkan kerangka kerja prediksi harga saham generik, dan menyematkan enam model berbeda dengan pendekatan analisis yang berbeda. Untuk melangkah lebih jauh, kami menggunakan kamus psikologis Harvard dan kamus sentimen keuangan Loughran-McDonald untuk membangun ruang sentimen. Artikel berita teks kemudian diukur secara kuantitatif dan diproyeksikan ke ruang sentimen. Metode pelabelan instan dibahas dan diuji secara ketat. Kami mengevaluasi akurasi prediksi model dan membandingkan kinerja mereka secara empiris pada berbagai tingkat klasifikasi pasar yang berbeda. Percobaan dilakukan pada harga Bursa Saham Hong Kong historis selama lima tahun dan artikel berita. Hasil menunjukkan bahwa (1) pada tingkat saham individu, sektor, dan indeks, model dengan analisis sentimen lebih unggul daripada model kata-kantong baik dalam set validasi maupun set pengujian independen; (2) model yang menggunakan polaritas sentimen tidak dapat memberikan prediksi yang berguna; (3) ada perbedaan minor antara model yang menggunakan dua kamus sentimen yang berbeda.

Kata kunci: Dampak berita, Perubahan harga saham, Sentimen, Prediksi, Eksperimen.

I PENDAHULUAN

1.1 Latar Belakang

Pasar saham menjadi salah satu elemen vital dalam sistem keuangan global, yang memiliki peran utama dalam alokasi modal dan pembentukan nilai (Raddant *et al.* 2021). Dalam beberapa dekade terakhir, terjadi peningkatan signifikan dalam keterbukaan politik dan ekonomi, yang sering diiringi dengan integrasi keuangan yang lebih luas. Namun, krisis keuangan dan ekonomi yang terjadi pada masa lalu menunjukkan bahwa ketidakstabilan dalam aktivitas keuangan dapat menyebabkan kontagion yang melampaui lingkup pasar keuangan itu sendiri. Oleh karena itu, penting untuk memahami interkoneksi antar pasar keuangan serta faktor-faktor yang mempengaruhinya, termasuk analisis berita dan informasi yang memengaruhi perilaku pasar.

Dalam beberapa aplikasi, makna yang terkait dengan konten sebuah gambar sering kali dipahami dalam hal distribusi spasial dari properti visual. Salah satu keterbatasan yang dikenal

dari model Bag of Words (BoW) adalah ketidakmampuannya untuk mengkodekan distribusi spasial kata visual dalam sebuah gambar (Silva *et al.* 2018). Oleh karena itu, kami memperkenalkan Pendekatan Bag of Visual Graphs (BoVG), yang menggunakan grafik untuk mengkodekan hubungan spasial antara pola visual ke dalam representasi gambar. Kelebihan pendekatan ini adalah kemampuannya untuk mengatasi keterbatasan model BoW dengan mempertimbangkan informasi spasial dalam representasi gambar, sehingga memungkinkan pengenalan pola visual yang lebih akurat dan komprehensif.

Analisis sentimen memodelkan dokumen dari dimensi sentimen. Alih-alih hanya mengukur frekuensi kata, setiap kata (terutama yang berwarna yang memiliki polaritas sentimen) diuraikan dan direpresentasikan oleh vektor fitur sentimen. Membuat prediksi berdasarkan representasi sentimen harus memiliki banyak keuntungan dibandingkan *bag-of-words*: (1) Dimensi reduktif (Kim 2018). Dibandingkan dengan puluhan ribu kata yang biasanya diamati saat menggunakan kata sebagai fitur, representasi sentimen sangat mengurangi dimensi menjadi ratusan, misalnya, kamus psikologi Harvard IV-4 memiliki 182 dimensi dan kamus sentimen keuangan Loughran-McDonald memiliki 6 dimensi; (2) Interpretasi afektif (Cambria *et al.* 2017). Biasanya sulit untuk menafsirkan pemetaan yang dihasilkan dalam ruang bag-of-words. Sebaliknya, skor sentimen di berbagai dimensi terkadang dapat memberikan orang perasaan yang lebih langsung tentang dokumen.

Meskipun analisis sentimen dokumen telah diusulkan dan digunakan dalam banyak aplikasi, misalnya, sistem rekomendasi (Bauman *et al.* 2017), belum banyak penelitian yang dilaporkan dalam domain silang ilmu komputer dan perdagangan algoritmik. Dalam penelitian ini, kami pertama kali membangun kerangka kerja generik yang dapat mengambil kumpulan informasi pasar dan memasukkan berbagai model prediksi khusus. Kami kemudian mengimplementasikan model prediksi menggunakan pendekatan analisis sentimen, pendekatan bag-of-words, atau pendekatan polaritas sentimen, dan membandingkan akurasi prediksi pengembalian harga saham harian mereka pada data pasar Hong Kong Stock Exchange selama lima tahun. Untuk menghindari bias yang diperkenalkan oleh kerangka kerja, seperti kamus sentimen, metode pelabelan, dan perbandingan pada tingkat klasifikasi pasar yang berbeda, kami (1) menggunakan dua kamus sentimen untuk perbandingan; (2) mendiskusikan dan menguji secara ketat berbagai metode pelabelan instance; dan (3) membandingkan kinerja model pada tingkat saham individu, sektor, dan indeks dalam eksperimen.

1.2 Tujuan

Penelitian ini bertujuan untuk mengembangkan model prediksi harga saham yang lebih akurat dengan mengintegrasikan analisis sentimen ke dalam kerangka kerja prediksi. Dengan memanfaatkan informasi sentimen dari artikel berita keuangan, model yang diusulkan diharapkan dapat menangkap hubungan antara sentimen pasar dan pergerakan harga saham dengan lebih baik dibandingkan pendekatan bag-of-words tradisional. Penelitian ini memberikan kontribusi penting dalam memahami peran analisis sentimen dalam prediksi pasar saham dan menawarkan solusi yang potensial untuk meningkatkan akurasi prediksi harga saham.

II TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen didefinisikan sebagai bidang studi yang menganalisis opini, sentimen, evaluasi, sikap, dan emosi melalui entitas dan hubungan mereka untuk entitas dan atribut terkait seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atribut mereka (Cambria *et al.* 2017). Tujuannya adalah untuk mengekstrak atribut dan komponen seperti opini, pengaruh, emosi, dan sentimen dalam teks dengan mempertimbangkan konteks semantik dan pragmatik.

2.2 Pendekatan Berbasis Kamus

Pendekatan ini memanfaatkan kamus sentimen yang terdiri dari kata-kata yang telah ditandai dengan polaritas atau dimensi sentimen. Setiap kata dalam teks dipetakan ke dalam kamus untuk menentukan sentimen atau orientasi sentimennya (Cambria *et al.* 2017).

Dua kamus sentimen yang digunakan adalah kamus sentimen Harvard IV-4 (HVD) dan kamus sentimen keuangan Loughran-McDonald (LMD), yang keduanya dibuat secara manual. Kamus-kamus manual diyakini lebih akurat dibandingkan yang dibuat secara otomatis karena pemilihan yang cermat oleh ahli linguistik. HVD memiliki lebih dari 10.000 kata dan mengkategorikan dimensi sentimen ke dalam 15 kelompok. Setiap kata dianalisis dan diproyeksikan ke 182 dimensi sentimen. LMD, yang dibuat oleh Loughran dan Bill McDonald pada tahun 2012, berisi lebih dari 3911 kata. Informasi detail mengenai kamus-kamus ini dapat ditemukan dalam literatur terkait.

2.3 Pendekatan Berbasis Konsep

Pendekatan ini mengeksplorasi pengetahuan konseptual seperti ConceptNet, ontologi, dan basis pengetahuan lainnya untuk melakukan analisis sentimen pada level konsep, bukan hanya pada level kata (Cambria *et al.*, 2017).

2.4 Deep Learning untuk Analisis Sentimen

Agarwal *et al.* (2020) memberikan penjelasan komprehensif tentang pendekatan deep learning seperti Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Recursive Neural Networks, dan pendekatan hybrid untuk analisis sentimen teks.

DAFTAR PUSTAKA

- Raddant, M., & Kenett, D. Y. 2021. Interconnectedness in the global financial market. *Journal of International Money and Finance*, 110, 102280.
- Silva, F. B., Werneck, R. D. O., Goldenstein, S., Tabbone, S., & Torres, R. D. S. 2018. Graph-based bag-of-words for classification. *Pattern Recognition*, 74, 266-285.

- Kim, K. 2018. An improved semi-supervised dimensionality reduction using feature weighting: application to sentiment analysis. *Expert Systems with Applications*, 109, 49-65.
- Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (Eds.). 2017. A practical guide to sentiment analysis (Vol. 5). Cham: Springer International Publishing.
- Bauman, K., Liu, B., & Tuzhilin, A. 2017. Aspect based recommendations: Recommending items with the most valuable aspects based on user reviews. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 717-725).
- Agarwal, B., Nayak, R., Mittal, N., & Patnaik, S. (Eds.). 2020. *Deep learning-based approaches for sentiment analysis* (Vol. 12, p. 319). Singapore: Springer.