

Critical Review 'Predicting the Future with Social Media'

Oleh

Cagiva Chaedar Bey Lirna – Accelerated Machine Learning

Sebuah *Journal Critical Review*, Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010*, 1, 492–499. <https://doi.org/10.1109/WI-IAT.2010.63>

Jurnal yang berjudul '*Predicting the Future with Social Media*' ini merupakan hasil penelitian yang dilakukan oleh Asur dan Huberman. Mereka melakukan penelitian ini dilatarbelakangi oleh ketertarikan mereka terhadap perkembangan sosial media sebagai wacana online yang mengandung keragaman informasi data yang sangat luar biasa. Menurut peneliti, dalam kasus media sosial, keragaman informasi yang menyebar melalui populasi pengguna media sosial yang besar menghadirkan peluang menarik untuk memanfaatkan penyebaran data tersebut ke dalam model prediksi tentang hasil tertentu. Dari informasi – informasi tersebut dapat dibangun model untuk mengumpulkan pendapat dari populasi pengguna media sosial secara kolektif dan mendapatkan wawasan baru yang berguna tentang berbagai perilaku yang timbul dalam populasi tersebut sembari memprediksi tren masa depan yang akan muncul dari populasi pengguna media sosial. Peneliti menambahkan, dengan kemudahan penggunaan, kecepatan, dan jangkauannya, media sosial seperti facebook, twitter, dll mampu dengan cepat mengubah wacana publik di masyarakat dan menetapkan tren dan agenda dalam topik yang jangkauannya sangat luas seperti politik, teknologi, dan hiburan (Asur & Huberman, 2010).

Berdasarkan pengembangan latar belakang yang disebutkan diatas, kemudian peneliti memutuskan bahwa penelitian dalam jurnal ini dilakukan dengan mengamati informasi yang tersebar di media sosial seraya melakukan perancangan model prediksi terhadap kemungkinan tren yang akan muncul di masa depan berdasarkan topik yang diangkat. Peneliti memfokuskan penelitiannya dengan memprediksi korelasi antara publisitas di *Twitter* dan implikasi kehidupan nyata terhadap pendapatan *box-office* terkait adanya penayangan film tertentu. Pada penelitian ini Asur dan Huberman mengelaborasi penelitian terdahulu milik (Gruhl et al., 2005; Joshi et al., 2010; Mishne & Glance, 2006; Sharda & Delen, 2006; Zhang & Skiena, 2009) kedalam penelitiannya. Dengan mengacu dan membandingkan pada penelitian terdahulu, Asur dan Huberman bertujuan menjalankan penelitiannya dengan mengidentifikasi 2 hal yaitu:

1. Bagaimana *buzz* dan perhatian terhadap suatu film dapat tercipta dan berubah seiring berjalannya waktu.
2. Apa faktor terciptanya *tweet* berupa sentimen positif dan negatif, serta bagaimana cara sentimen tersebut menyebar dan memengaruhi perilaku populasi pengguna *Twitter*.

Adapun dataset *tweet* yang digunakan peneliti diperoleh dengan *crawling* data tweet di Twitter. Hasil *crawling* didapat sebanyak 2.89 juta *tweet* yang mengacu pada 24 film yang berbeda. Kemudian data *tweet* tersebut disaring kembali dengan hanya menganalisis data yang *keyword*-nya berupa film yang akan segera tayang. Setelah data penelitian berhasil didapat, kemudian peneliti merumuskan model regresi yang sesuai untuk memprediksi pendapatan *box-office* di dunia nyata berdasarkan pengaruh sentimen dari publisitas pengguna *twitter* mengenai film yang akan segera tayang.

Hasil jurnal menyimpulkan model regresi linier yang dirancang peneliti berhasil menunjukkan bahwa jumlah publisitas di media sosial berkorelasi dengan implikasi kehidupan nyata dalam memprediksi pendapatan *box-office*. Asur dan Huberman juga menyatakan bahwa

model yang mereka rancang dapat digeneralisasi untuk memprediksi topik pendapatan lain dan mereka berspekulasi bahwa modelnya dapat digeneralisasi lebih lanjut untuk memprediksi topik yang berbasis non-pendapatan.

Secara umum, penulisan penelitian ini sudah baik, dapat dilihat dari pemilihan kata yang digunakan penulis sangat mudah dipahami dan sudah mengikuti kaidah - kaidah dalam penyusunan penelitian ilmiah. Penyampaian setiap pokok bahasan yang dituliskan secara rapi, juga sangat membantu pembaca untuk mengetahui apa yang ingin disampaikan peneliti. Hasil penelitian ini juga memberi wawasan baru bagi peneliti lain yang ingin mengembangkan model prediksi yang berhasil dirumuskan dalam jurnal. Sayangnya, perumusan model prediksi dituliskan peneliti dengan tahapan yang rumit dan kurang jelas, sehingga cukup susah untuk memahami bagaimana perumusan model prediksi berhasil dibuat. Tak hanya itu, pada proses pengolahan data sampel dari *crawl* hingga penyaringan data terkesan rumit dan terlalu banyak grafik yang membingungkan, sehingga sulit untuk memahami bagaimana peneliti memperoleh data uji final.

Singkatnya, dapat dikatakan penelitian saat ini masih jauh dari konklusif dan masih banyak hal yang harus digarisbawahi sebagai bahan evaluasi. Langkah-langkah yang lebih baik harus dikembangkan, dan proses pengolahan data sampel dan data uji harus diperjelas untuk meningkatkan pemahaman kita terkait pemanfaatan data uji terhadap proses perumusan model prediksi. Terlepas dari beberapa kekurangan dalam metodologi, studi lebih lanjut perlu dilakukan untuk mencoba menyelidiki topik lain di media sosial dengan target capaian yang lebih luas dari penelitian yang sudah dilakukan oleh Asur dan Huberman ini.

Referensi

- Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010*, 1, 492–499. <https://doi.org/10.1109/WI-IAT.2010.63>
- Gruhl, D., Guha, R., Kumar, R., Novak, J., & Tomkins, A. (2005). The Predictive Power of Online Chatter. *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, 78–87. <https://doi.org/10.1145/1081870.1081883>
- Joshi, M., Das, D., Gimpel, K., & Smith, N. A. (2010). Movie Reviews and Revenues: An Experiment in Text Regression. *Human Language Technologies: The 2010 Annual Conference of the North {A}merican Chapter of the Association for Computational Linguistics*, 293–296. <https://aclanthology.org/N10-1038>
- Mishne, G., & Glance, N. (2006). Predicting movie sales from blogger sentiment. *AAAI Spring Symposium - Technical Report, SS-06-03*, 155–158.
- Sharda, R., & Delen, D. (2006). Predicting box-office success of motion pictures with neural networks. *Expert Systems with Applications*, 30(2), 243–254. <https://doi.org/10.1016/j.eswa.2005.07.018>
- Zhang, W., & Skiena, S. (2009). Improving movie gross prediction through news analysis. *Proceedings - 2009 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2009*, 1, 301–304. <https://doi.org/10.1109/WI-IAT.2009.53>