



Analisis Sentimen Opini Masyarakat Terhadap Bakal Calon Walikota Surabaya 2020 Berdasarkan *Social Media Mining* Menggunakan Algoritma *N-Gram-Multichannel CNN*



Ferisa Tri Putri Prestasi, Mohammad Isa Irawan
Departemen Matematika Institut Teknologi Sepuluh Nopember

Abstrak

Pilwali Surabaya mendapat menjadi perhatian publik, dimana para bakal calon mulai mengajukan diri dan ramai diperbincangkan di media sosial, diantaranya Facebook dan Twitter. Penting bagi bakal calon untuk mengetahui sentimen opini yang berkembang di media sosial dengan menggunakan implementasi analisa sentimen pada *social media mining* untuk dapat di klasifikasi sentimen positif atau negatif dari opini tersebut. Oleh karena itu, pada penelitian analisa sentimen ini digunakan algoritma *N-Gram Multichannel CNN* untuk mendapatkan analisa sentimen dengan memanfaatkan konsep *Natural Language Processing* yang memungkinkan komputer untuk memproses dan memahami bahasa alami manusia dan memperoleh akurasi model optimal terhadap dataset opini. Hasil diperoleh analisis sentimen pada teks berisikan opini masyarakat dari media sosial terhadap bakal calon Walikota Surabaya 2020 dapat diterapkan dengan baik menggunakan Algoritma *N-Gram-Multichannel CNN* dengan diperoleh akurasi model terhadap data latih sebesar 94.38% dan 96.12% pada data uji.

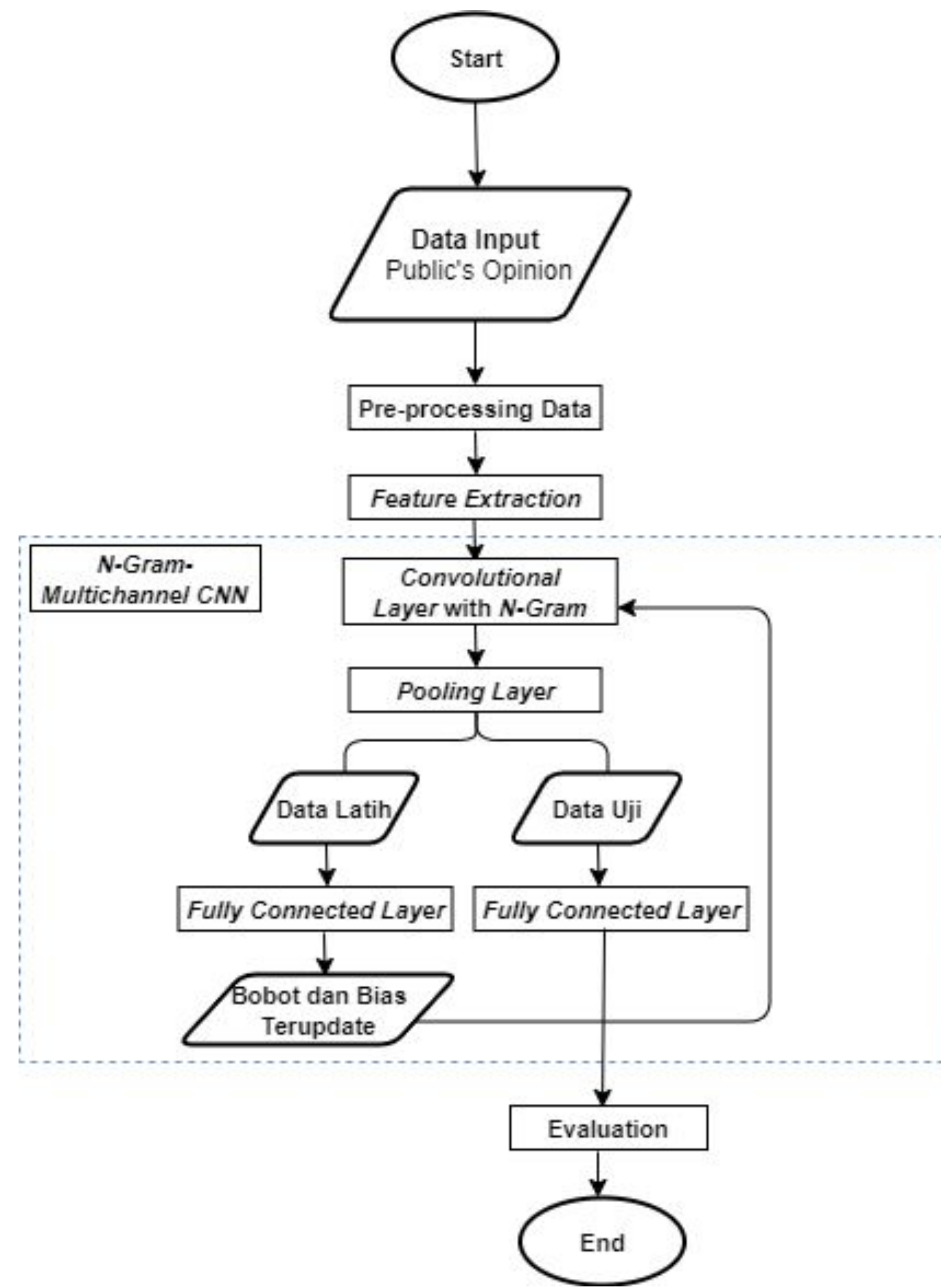
Kata Kunci: Pilwali Surabaya, Facebook, Twitter, *Social Media Mining*, Analisis Sentimen, *Natural Language Processing*, *N-Gram-Multichannel CNN*

Latar Belakang

Indonesia adalah salah satu negara yang menganut sistem demokrasi . Hal ini ditandai dengan diadakannya suatu pemilihan umum terhadap Walikota dan Wakil Walikota salah satunya di kota Surabaya. Pemilihan umum pada suatu negara yang menganut sistem demokrasi diselenggarakan secara periodik. Komisi Pemilihan Umum (KPU) RI telah menetapkan jadwal Pemilihan Wali Kota (Pilwali) Surabaya yang akan datang. Bahkan pemilihan kepala daerah (Pilkada) itu akan dilaksanakan secara serentak di seluruh Indonesia. Seorang tokoh politik yang ingin maju sebagai bakal calon Walikota Surabaya akan melihat atau mempertimbangkan popularitas mereka berdasarkan opini masyarakat. Kemeriahan Pilwali Surabaya 2020 sudah dirasakan di media sosial khususnya Twitter dan Facebook yang sekarang ini menjadi tempat yang sangat penting untuk calon dan tim suksesnya melakukan kampanye. Hal ini memicu berbagai opini di Twitter. Opini publik memiliki peran penting dalam menyukseskan bakal calon dalam pemilu, bahwa opini publik dapat memberikan pengaruh kepada bakal calon dalam menentukan sikap. Opini ini dapat dimanfaatkan untuk melihat bagaimana polaritas tokoh politik yang akan maju sebagai bakal calon Walikota dan Wakil Walikota Surabaya tahun 2021. Penentuan polaritas positif atau negatifnya suatu opini dapat dilakukan secara manual, tetapi seiring bertambahnya sumber opini yang semakin banyak, tentunya waktu dan usaha yang dibutuhkan untuk mengklasifikasikan polaritas opini tersebut akan semakin banyak. Pada penelitian ini, analisis sentimen dilakukan untuk melihat dan mengambil informasi berupa opini seseorang dalam bahasa indonesia di Twitter dan Facebook yang ditujukan kepada bakal calon Walikota dan Wakil Walikota Surabaya periode 2021-2024, apakah opini itu masuk kategori opini positif atau negatif dengan menggunakan kombinasi Algoritma *N-Gram-Multichannel CNN* dalam pengklasifikasian suatu teks yang dievaluasi hingga menghasilkan akurasi model paling optimal.

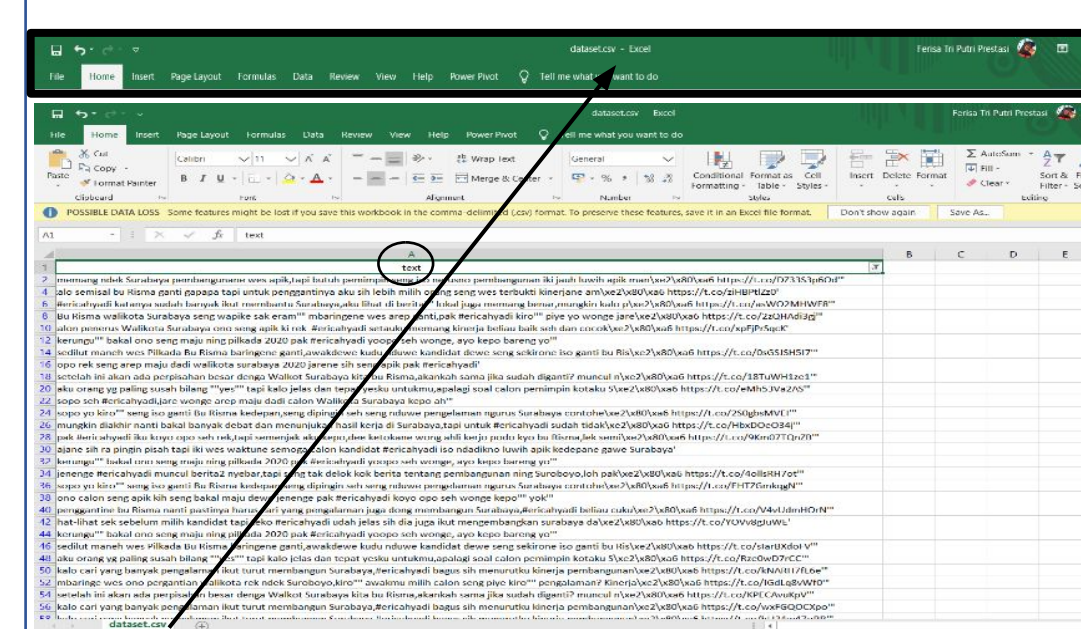
Metodologi

Blok Diagram metode penelitian yang dilakukan untuk menyelesaikan permasalahan ini tampak pada Gambar dibawah ini:



Gambar 1. Metodologi Penelitian

Perancangan dan Implementasi



Gambar 2. Input Kalimat Opini File CSV

jaman sekarang jangan asal memilih ya cari yg pengalaman dan sudah banyak kinerja yang bagus untuk bisa diyakini bisa maju,erichayadi menurutku juga sudah jelas dan bakal ga salah memilihnya

Gambar 3. Hapus URL

jaman sekarang jangan asal memilih ya cari yg pengalaman dan sudah banyak kinerja yang bagus untuk bisa diyakini bisa maju erichayadi menurutku juga sudah jelas dan bakal ga salah memilihnya

Gambar 4. Hapus Tanda Baca, Simbol, dan Angka

jaman sekarang jangan asal memilih ya cari yg pengalaman dan sudah banyak kinerja yang bagus untuk bisa diyakini bisa maju erichayadi menurutku juga sudah jelas dan bakal ga salah memilihnya

Gambar 6. Casefolding

jaman sekarang jangan asal memilih ya cari yg pengalaman dan sudah banyak kinerja yang bagus untuk bisa diyakini bisa maju erichayadi menurutku juga sudah jelas dan bakal ga salah memilihnya

Gambar 5. Hapus Duplikat Huruf

[jaman,'sekarang','jangan','asal','memilih','ya','cari','yg','pengalaman','dan','sudah','banyak','ki','nerja','yang','bagus','untuk','bisa','diyakini','bisa','maju','erichayadi','menurutku','juga','sudah','je','las','dan','bakal','ga','salah','memilihnya']

Gambar 7. Tokenisasi

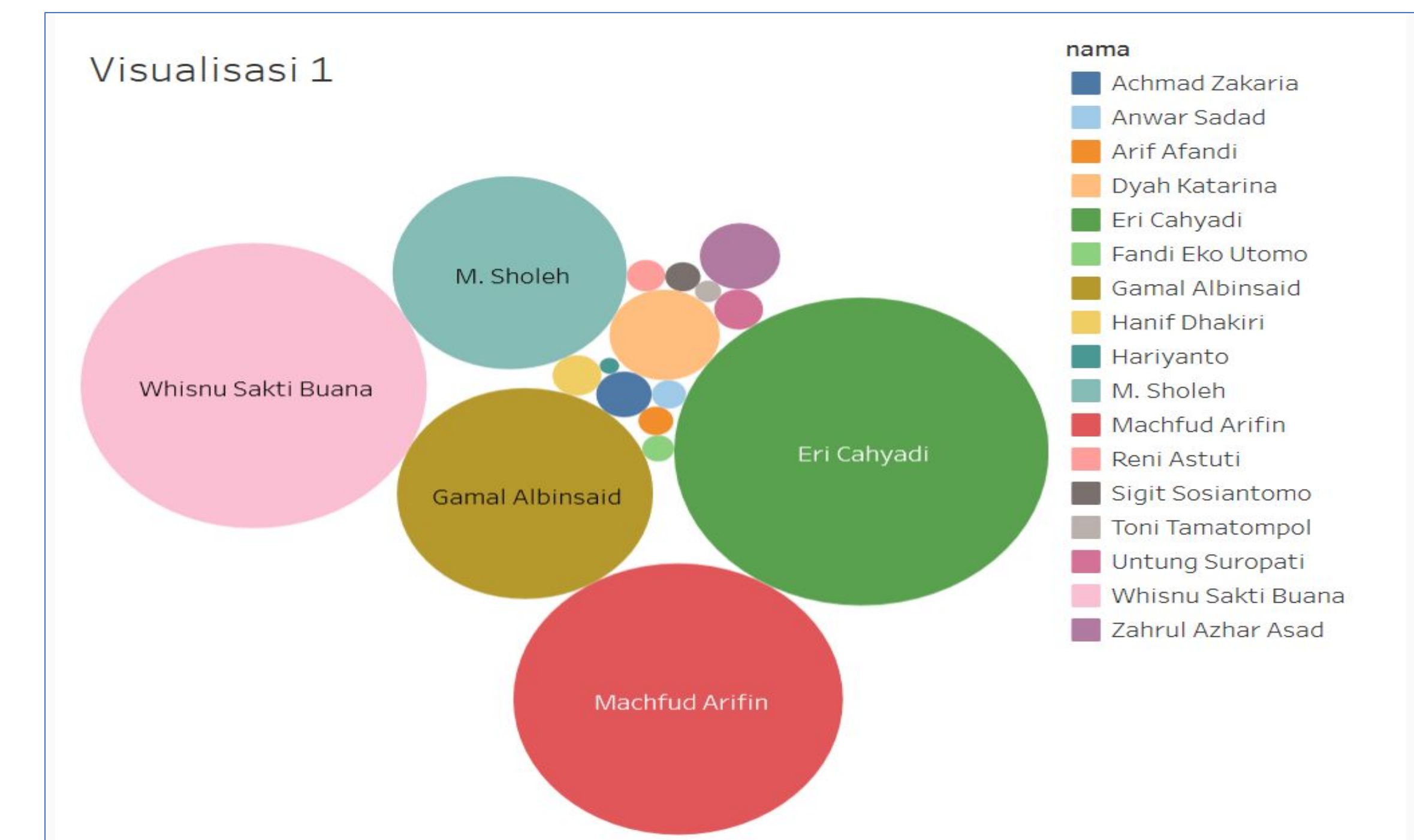
[jaman,'sekarang','jangan','asal','memilih','ya','cari','yg','pengalaman','dan','sudah','banyak','ki','nerja','yang','bagus','untuk','bisa','diyakini','bisa','maju','erichayadi','menurutku','juga','sudah','je','las','dan','bakal','nggak','salah','memilihnya']

Gambar 8. Spelling Norm

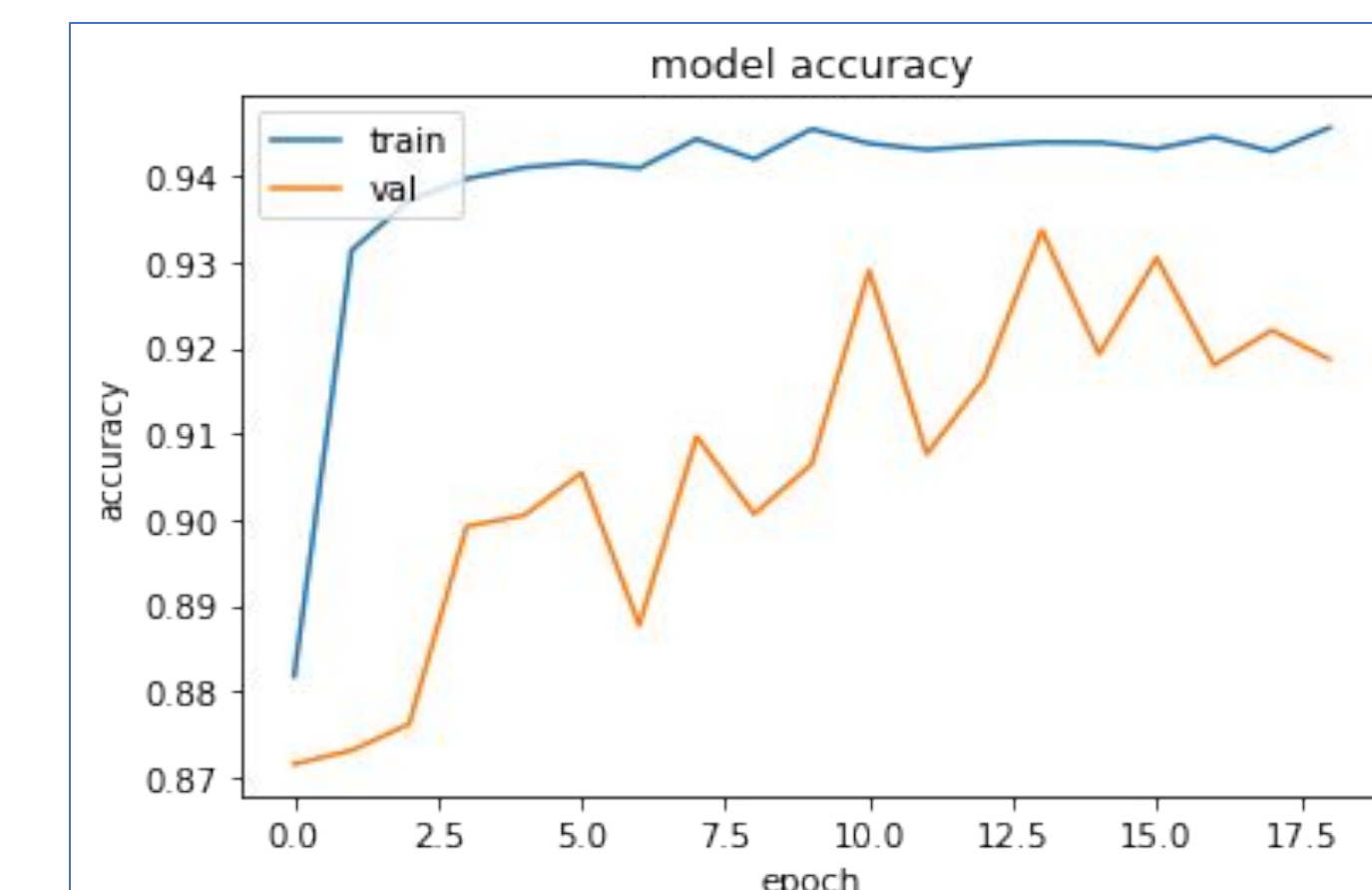
[jaman,'sekarang','asal','memilih','cari','pengalaman','sudah','banyak','kinerja','bagus','bisa','diyakini','maju','erichayadi','menurutku','bakal','nggak','salah','memilihnya']

Gambar 9. Filtering

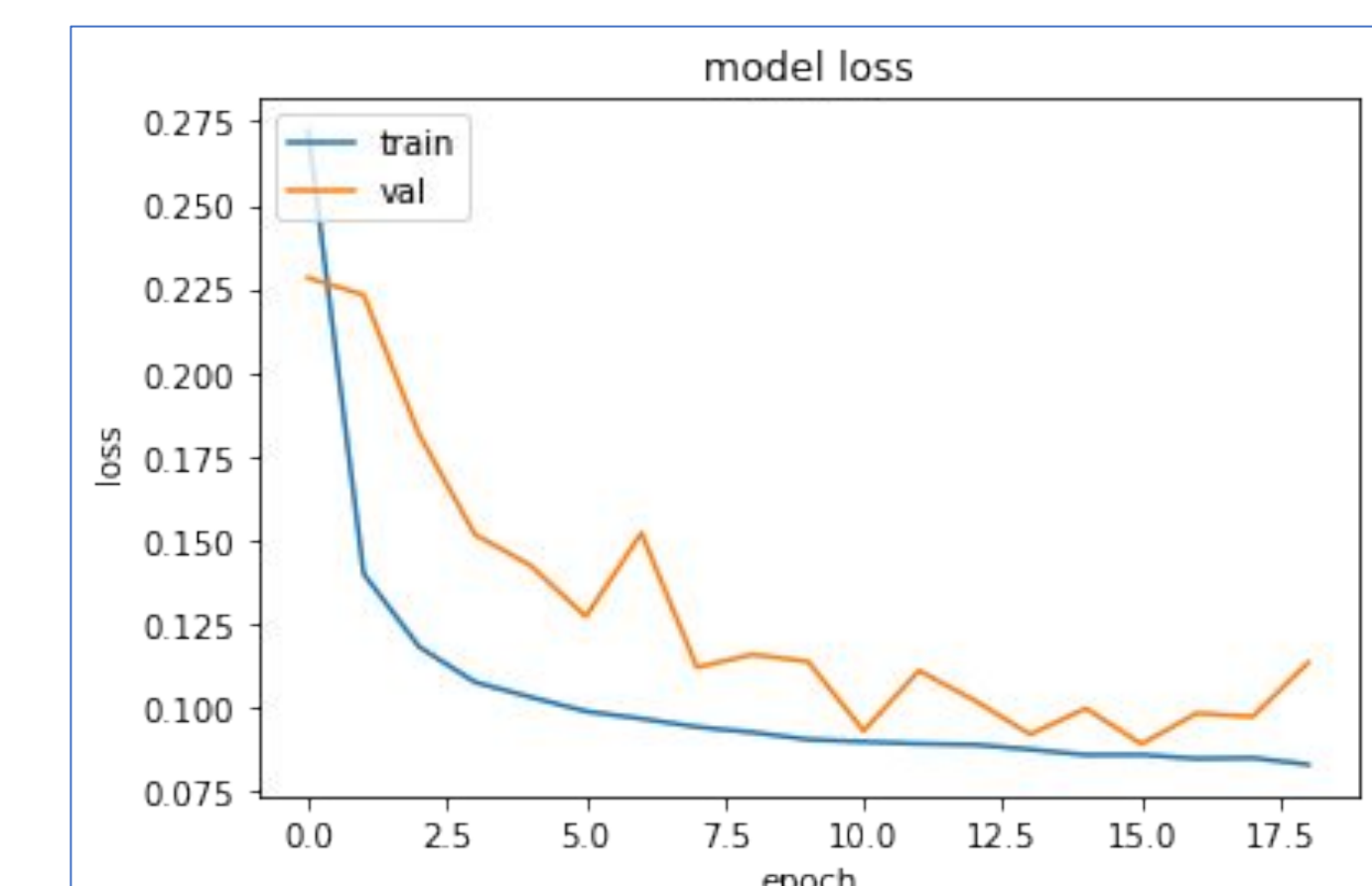
Hasil



Gambar 10. Visualisasi Data Jumlah Opini Masing-masing Nama Bakal Calon Walikota Surabaya 2020



Gambar 11. Grafik akurasi model terhadap data latih dan data uji



Gambar 12. Grafik pergerakan loss model

Kesimpulan

Cara menerapkan analisis sentimen opini masyarakat terhadap bakal calon walikota Surabaya 2020 yaitu dengan melakukan *crawling* data, pra-pemrosesan data yang meliputi penghapusan url, penghapusan tanda, penghapusan huruf yang berulang, *case folding*, tokenisasi, *spelling normalization*, dan filterisasi, selanjutnya pelabelan data, ekstraksi fitur, dan yang terakhir pembentukan model klasifikasi menggunakan algoritma *N-Gram-Multichannel CNN* meliputi konfigurasi *Hyperparameter* pada layer konvolusi dan *embedding*, *Max Pooling*, *Flatten*, dan *Dense* dengan mengatur fungsi aktivasi. Tahapan analisis sentimen ini dapat diterapkan dengan baik yang menghasilkan data pelabelan didapatkan 8.312 dataset negatif dan 34.579 dataset positif. Kata yang sering muncul pada dataset Negatif adalah “salah” sebanyak 18.620 kata, “pencitraan” sebanyak 6.023 kata, “jelek” sebanyak 5.879 kata. Pada dataset Positif terdapat kata “bagus” sebanyak 31.418 kata, “cocok” sebanyak 20.800 kata, dan “terbukti” sebanyak 9.679 kata. Selain itu, diperoleh hasil akurasi model terhadap data latih dan data uji dataset opini masyarakat berdasarkan Social Media Mining terhadap bakal calon Walikota Surabaya 2020 dalam menganalisis sentimen menggunakan Algoritma *N-Gram-Multichannel CNN* diperoleh akurasi model terhadap data latih sebesar 94.38% untuk data latih dan sebesar 96.12% untuk data uji dengan *Learning Curve* yang menunjukkan *Good fit* setelah dilakukan *Tuning Parameter*.

Kontak

Ferisa Tri Putri Prestasi
Departemen Matematika FSAD
Institut Teknologi Sepuluh Nopember
Email : ferisatri04@gmail.com
Github: <https://github.com/prestasicode>
Medium : <https://medium.com/@ferisaprestasi>



References

1. S. W. A. D. F. N. Sihwi, “Analisa Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier,” Research Gate Publication, 2016.
2. R. Zafarani, M. A. Abbasi and H. Liu, *Social Media Mining : An Introduction*, London: Cambridge University Press, 2016.
3. R. W. M. L. S. N. H. A. M. R. M. I. Irawan, “Implementation of social media mining for decision making in product,” Journal of Physics: Conference Series, p. Conf. Ser. 1490 012068, 2020.
4. B. Y. R. F. R. J. D. T. R. M. Y. F. Haryanto, “Facebook Analysis of Community Sentiment on 2019 Indonesian Presidential Candidates from Facebook Opinion Data,” ScienceDirect Procedia Computer Science The Fifth Information System International Conference 2019., 2019.
5. G. A. Buntoro, “Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter,” Researchgate. INTEGER: Journal of Information Technology 2.1, 2019.
6. G. Asrofi, “Analisis Sentimen Calon Gubernur Jawa Timur 2018 dengan Metode Naive Bayes Classifier,” Journal of Informatics Pelita Nusantara, 2019.
7. R. Ardiansyah, “Analisis Sentimen Calon Presiden dan Wakil Presiden Periode 2019-2024 Pasca Debat Pilpres di Twitter,” ScientiCO - Computer Science and Informatics Journal, vol. 2, no. 1, pp. E-ISSN: 2620-4118, 2019.
8. T. D. S. S. Y. S. Lukmana, “Analisis Sentimen Terhadap Calon Presiden 2019 dengan Support Vector Machine di Twitter,” Seminar Nasional Penelitian Pendidikan Matematika (SNP2M) 2019 UMT, 2019.
9. I. E. Firdausi, “Analisis Sentimen Tanggapan Pelanggan Operator Telekomunikasi di Twitter dengan Algoritma DCNN-SVM,” 2019.
10. M. K. M. Rozi, “Opinion mining on book review using CNN-L2-SVM algorithm,” Journal of Physics: Conference Series. Pg. 012004., vol. 974, p. 012004, 2018.