

Penerapan Algoritma Naïve Bayes Untuk Analisis Sentimen Pengguna My JNE

Regina Aswini Puspala Sari¹, Irfan Dwi Jaya²

¹ UIN Raden Fatah Palembang and sariregina90@gmail.com

² UIN Raden Fatah Palembang

ABSTRAK

Teknologi diketahui dapat mengubah gaya hidup manusia secara fundamental, salah satunya dalam aspek bisnis. Perusahaan JNE sebagai salah satu perusahaan ekspedisi tentunya menerapkan langkah untuk mengetahui bagaimana opini masyarakat terhadap pelayanan yang diberikan oleh perusahaan melalui sebuah aplikasi My JNE di google store. Hal ini bertujuan agar pihak perusahaan dapat mengetahui ulasan positif maupun ulasan negatif dari aplikasi tersebut sehingga dapat menjadi bahan pertimbangan evaluasi bagi perusahaan. Saat melakukan analisis perasaan baik dan negatif, penggunaan algoritma *Naïve bayes* diperlukan karena kebutuhan ini. Analisis mengungkapkan bahwa total 1.665 ulasan pelanggan JNE dikumpulkan dari situs web Google Play antara tanggal 2 Juni 2021 dan 30 September 2021. Jumlah pengguna terbesar, 1.123, telah memberi item tersebut peringkat 1, yang mana sesuai dengan persentase 67,45%. Sedangkan pada pelabelan kelas opini diketahui terdapat 1.603 review yang mampu diproses mengikuti tahap preprocessing, yang kemudian diperoleh review negatif sebanyak 1.285 review (80,16%) dan review positif sebanyak 318 review (19,84%).

Kata Kunci: *Naïve bayes, sentimen, JNE.*

ABSTRACT

Technology is known to fundamentally change human lifestyles, one of which is in the business aspect. The JNE company as one of the shipping companies certainly implements steps to find out how the public thinks about the services provided by the company through a My JNE application on the Google Store. This is intended so that parties can find out positive reviews and negative reviews of the application so that it can be used as material for evaluation considerations for the company. Therefore, the use of the naïve Bayes algorithm is necessary in analyzing positive and negative sentiments. The classification results show that for the period 02 June 2021 to 30 September 2021, 1,665 reviews were obtained from JNE users on the Google Play website. While most users give a rating of 1 with a proportion of 67.45% or as many as 1,123 users. Whereas in the sentiment class labeling after preprocessing it was found that there were 1,603 reviews that could be processed, which then obtained a total of 1,285 negative reviews (80.16%) and positive reviews of 318 reviews (19.84%).

Keywords: *Naïve bayes, sentimen, JNE.*

PENDAHULUAN

Sudah menjadi rahasia umum bahwa kemajuan teknologi dapat mengubah gaya hidup manusia secara drastis, salah satunya adalah dunia komersial. Penggunaan teknologi informasi adalah salah satu strategi yang digunakan bisnis untuk memenangkan klien baru. Karena semakin banyak orang yang menggunakan internet, berbagai bisnis kini bersaing satu sama lain untuk menjual produk mereka, baik barang maupun jasa, menggunakan media internet. Menurut hasil survei yang dilakukan oleh APJII (2018), diketahui bahwa salah satu dari 10 alasan teratas responden di Indonesia memanfaatkan internet adalah untuk melakukan penjualan online. Meskipun melakukan pembelian secara online adalah salah satu dari 15 alasan utama penggunaan internet, ada banyak keuntungan lainnya juga. E-commerce, yang secara sederhana mengacu pada praktik menjalankan bisnis secara online, telah menghasilkan menjamurnya beberapa pasar online. Selain itu, Chrisbiyanto (2018) menegaskan bahwa ekspansi e-commerce tentunya didorong oleh bisnis

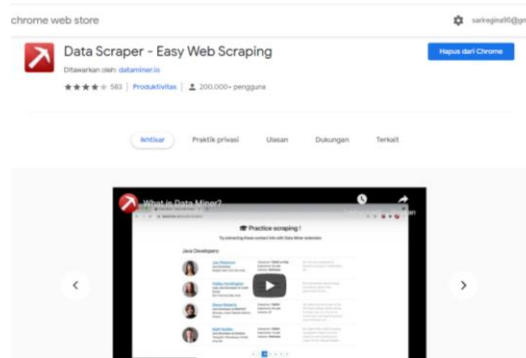
logistik, khususnya perusahaan pelayaran. E-commerce menjadi lebih mudah, dari pergudangan hingga distribusi, dengan munculnya layanan pengiriman barang. Akibatnya, e-commerce dapat lebih berkonsentrasi pada penjualan tanpa harus khawatir tentang bagaimana barang akan diangkut.

Menjamurnya bisnis yang menyediakan layanan pengiriman barang merupakan indikasi lain bahwa industri logistik semakin matang sebagai sebuah industri. PT. Tiki Jalan Nugraha Ekakurir adalah salah satu perusahaan jasa pengiriman dan logistik (JNE) yang cukup terkenal dan banyak digunakan. Mengingat informasi yang ditunjukkan di atas, penting untuk melakukan penelitian lebih lanjut untuk memastikan pendapat pengguna tentang aplikasi MyJNE. Penelitian harus dilakukan dengan tujuan akhir untuk memberikan umpan balik positif dan negatif kepada bisnis terkait perangkat lunak. Kemungkinan hal ini akan dinilai dan diperhitungkan di masa mendatang untuk memastikan bahwa konsumen puas dengan banyaknya fungsi yang disediakan. Akibatnya, strategi scraping digunakan untuk tujuan pengumpulan data karena strategi ini terbukti efektif dan efisien dalam hal biaya, waktu, dan tenaga. Review tersebut kemudian disortir ke dalam kategori positif dan negatif menggunakan metode *Naïve bayes Classifier*.

METODE PENELITIAN

Metode penelitian kualitatif deskriptif digunakan dalam penelitian ini sebagai jenis penelitian yang dilakukan. Sebuah metode yang dikenal sebagai penelitian kualitatif deskriptif adalah salah satu yang lebih menekankan pada mendapatkan pemahaman yang lebih mendalam tentang suatu masalah daripada hanya mengamati suatu masalah. (Gunawan, 2017) Metode penelitian kualitatif deskriptif dapat diartikan juga sebagai metode penelitian yang berlandaskan pada riset yang sifatnya deskripsi, cenderung menggunakan analisis dan lebih menampakkan proses maknanya. Tujuan dari metode ini adalah untuk memahami secara luas dan mendalam terhadap suatu masalah secara detail pada suatu permasalahan yang sedang dikaji. (Semaiwan, 2014)

Untuk keperluan investigasi ini, dikumpulkan data berupa review pengguna JNE dengan memanfaatkan aplikasi my JNE yang terdapat di website Google Play. Metode scraping digunakan untuk mengambil data, dan dilakukan dengan menggunakan ekstensi Google Chrome. Dimungkinkan untuk mengambil data dari banyak situs internet dengan menggunakan salah satu dari banyak ekstensi yang tersedia untuk Google Chrome. Data Scraper adalah salah satu program tersebut. Ekstraksi data gratis dari situs web dimungkinkan dengan menggunakan alat yang disebut pengikis data. Tentang penelitian ini, perangkat lunak Data Scraper, versi 3.299.84, digunakan. Data yang diekstraksi dapat diimpor ke spreadsheet yang dibuat dalam format Microsoft Excel atau CSV.

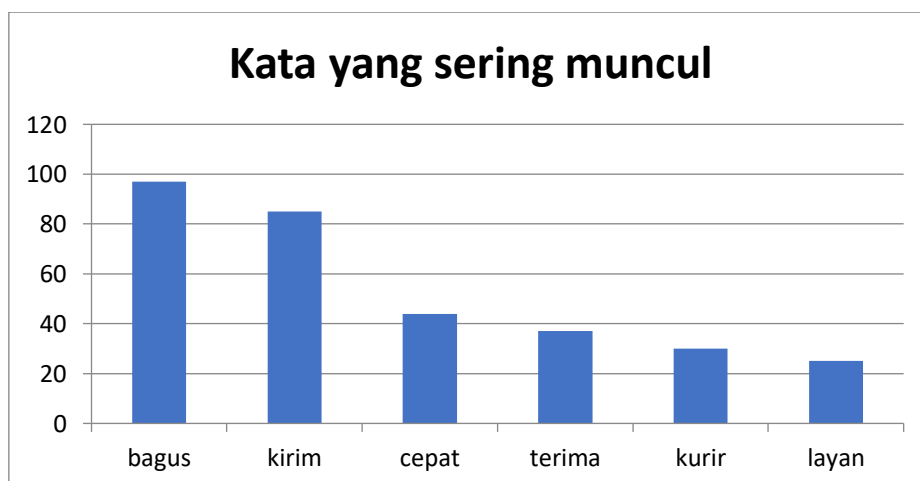


Gambar 1. Ekstensi Data Scraper

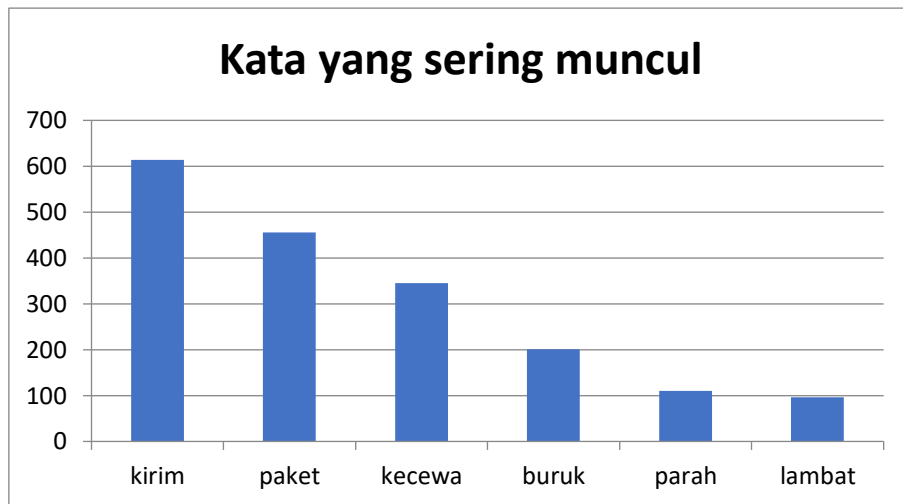
HASIL DAN PEMBAHASAN

Angka 1.665 ditetapkan sebagai total review yang dikumpulkan dari pengguna JNE selama tanggal 2 Juni 2021 dan 30 September 2021. Seperti terlihat pada Gambar 4.3, bulan September 2021 menjadi bulan dengan total review terbanyak. jumlah ulasan selama periode waktu ini. Kemudian terkait rating, sebanyak 1.123 pengguna memberikan rating 1 yang artinya tidak bagus. Setelah dilakukan studi deskriptif untuk mendapatkan gambaran secara keseluruhan tentang pengguna JNE, data selanjutnya dilakukan preprocessing. Data review yang awalnya berjumlah 1.665 review, dikurangi menjadi 1.603 review sebelum dikirim ke tahap pemrosesan selanjutnya. Karena evaluasi yang dipermasalahkan hanya terdiri dari emotikon atau simbol, bukan teks sebenarnya, evaluasi tersebut tidak dapat diselamatkan setelah data dibersihkan, yang mengakibatkan penurunan jumlah total ulasan secara signifikan. Ulasan tersebut awalnya dikategorikan berdasarkan tingkat kepositifannya, dengan 161 ulasan positif, 528 ulasan netral, dan 914 ulasan buruk yang membentuk tiga kelompok berbeda. Sejumlah besar kata-kata yang belum ditemukan serta kata-kata positif dan negatif dalam jumlah yang sama cenderung menjadi penyebab tingginya kuantitas sentimen netral. Pemberian label netral terhadap sentimen positif atau negatif tergantung dari beberapa pertimbangan karena sikap netral diyakini tidak memberikan masukan atau keuntungan apapun bagi JNE. Pelabelan kelas sentimen netral akan dilakukan secara manual, dengan ketentuan sebagai berikut: jika ulasan tidak mengidentifikasi kata-kata positif atau negatif, maka akan diklasifikasikan dalam kelas positif; jika jumlah kata positif atau negatif sama, maka akan diklasifikasikan sebagai kelas negatif dengan pertimbangan agar pihak yang berkepentingan dapat melihat permasalahan yang terjadi; dan jika ulasan tidak mengidentifikasi kata sama sekali, maka itu akan diklasifikasikan dalam kelas negatif.

Pada evaluasi yang berisi pernyataan positif seperti ungkapan pujian, ucapan terima kasih dan sebagainya termasuk dalam kategori sentimen positif. Kategori ulasan yang berisi sentimen negatif mencakup ulasan yang menampilkan pernyataan negatif seperti hinaan, ketidaksenangan, dan pernyataan serupa lainnya. Akibatnya, hanya ada dua jenis label yang digunakan, yaitu perasaan positif dan negatif. Menurut temuan pelabelan akhir, ditemukan bahwa jumlah ulasan yang mengungkapkan sikap negatif lebih sering terjadi daripada ulasan yang mengungkapkan suasana hati yang baik. Hal ini ditunjukkan pada Gambar 4.6, dimana ulasan pada kelas sentimen negatif berjumlah 80,16% atau 1.285 ulasan. Di sisi lain, 19,84% dari evaluasi, atau 318 di antaranya, menyampaikan perasaan yang baik.



Gambar 1. Kata yang sering muncul pada ulasan positif



Gambar 1. Kata yang sering muncul pada ulasan negatif

Klasifikasi Naïve Bayes

Memuat uraian tentang pembahasan hasil pengabdian kepada masyarakat dari awal proses pengabdian hingga perubahan masyarakat yang dihasilkan dari pengabdian kepada masyarakat. Pembahasan tentang hasil pengabdian kepada masyarakat didukung oleh referensi dan perspektif teoritis yang didukung oleh kajian literatur yang relevan. Referensi harus dilakukan mengikuti gaya APA 6th. (Palatino Linotype, ukuran 12, Spasi baris: 1,15 sebelum 6 poin, 1,15 setelah 6 poin)

Langkah pertama dalam proses klasifikasi menggunakan *Naïve bayes* Classifier adalah membagi jumlah total data menjadi dua kategori: data uji dan data pelatihan. Dalam ruang lingkup penyelidikan ini, perbandingan dilakukan lima kali dengan proporsi data latih dan data uji yang berbeda. Distribusi awal data dimulai dengan enam puluh persen menjadi data pelatihan (962 ulasan), dan empat puluh persen menjadi data uji (641 ulasan). Diketahui Presisi Kelas untuk data negatif sebesar 90,0%, tetapi Presisi Kelas untuk data positif sebesar 65,5%. Hal ini didasarkan pada sebaran data latih dan data uji. Akurasi yang dicapai adalah 81,3% sebagai Konsekuensinya.

Pembagian kedua data dilakukan dengan 70 persen data pelatihan (1.121 ulasan) dan 30 persen data uji (480 ulasan). Data latih dan data uji distribusi kedua menghasilkan akurasi yang lebih tinggi dari data distribusi pertama yaitu sebesar 83,65%. Kelas Presisi untuk kelas negatif dihasilkan sebesar 90,6% dari data latih dan data uji divisi II, sedangkan untuk kelas positif sebesar 70,4%. Pembagian akhir data dilakukan dengan 80% data menjadi data pelatihan (1.282 ulasan), dan 20% menjadi data uji (321 ulasan). Akurasi yang dicapai hasil pembagian ketiga data latih dan data uji lebih tinggi dibandingkan dengan akurasi yang dicapai pembagian data pertama dan kedua yaitu sebesar 85,67%. Kelas presisi Kelas negatif dihasilkan sebesar 89,7% dari waktu selama pembagian ini, sedangkan akurasi kelas positif meningkat menjadi 77,5%.

Pembagian keempat data dilakukan dengan 85% data berasal dari training (1.362 review), dan 15% berasal dari testing (241 review). Di divisi keempat data pelatihan dan data uji, 91,1% instans menghasilkan Kelas Presisi negatif, sementara persentase instan yang menghasilkan Kelas Presisi positif meningkat menjadi 79,5%. Hasilnya, akurasi yang dicapai lebih tinggi daripada tiga perbandingan data sebelumnya, yakni 87,25%.

Pemisahan data yang lengkap terdiri dari 10% data pengujian dan 90% data pelatihan (1.442 ulasan) (161 ulasan). Pada pendistribusian data latih dan data uji, yang terakhir menghasilkan Kelas

Presisi negatif sebesar 91,3%, sedangkan untuk kelas positif meningkat menjadi lebih dari 80% yaitu 81,5%, dengan akurasi yang dihasilkan lebih tinggi dari semua data sebelumnya. distribusi yaitu 87,97%.

Ketika metode *Naïve bayes Classifier* diterapkan, hasil klasifikasi menunjukkan bahwa berbagai data uji dan data latih menghasilkan akurasi yang cukup baik karena lebih besar dari 80%. Ini ditentukan berdasarkan hasil klasifikasi. Berkenaan dengan lima cara berbeda di mana data telah didistribusikan sebelumnya, perbandingan dilakukan sehingga dapat ditetapkan bahwa distribusi 90% data pelatihan dan 10% data uji adalah yang terbaik. Hal ini dibuktikan dengan nilai akurasi yang paling tinggi dibandingkan pembandingan lainnya yaitu 87,97% dengan Class Precision di atas 80%, dengan 91,3% untuk kelas negatif dan 81,5% untuk kelas positif.

Setelah itu ditetapkan, berdasarkan temuan klasifikasi *Naïve bayes*, bahwa selain akurasi, presisi dan daya ingat juga meningkat secara umum, yang disertai dengan pengurangan jumlah data uji yang digunakan. Kamus leksikal atau kumpulan kata digunakan untuk analisis sentimen dalam proyek penelitian khusus ini. Dalam metode yang disebut tas kata-kata, sistem menganalisis kata-kata yang ada dalam suatu sumber tanpa mengklasifikasikan kata-kata tersebut ke dalam kategori yang berbeda tergantung pada struktur kalimatnya, seperti kata benda, kata kerja, kata sifat, dan sebagainya. Hal inilah yang menyebabkan munculnya kesalahan dalam sistem pelabelan, yang disebabkan oleh kesalahpahaman kata-kata yang identik satu sama lain. Misalnya, mirip dengan frasa "tidak baik", istilah ini dapat diubah menjadi frasa "baik", meskipun sudah menjadi rahasia umum bahwa kedua frasa tersebut saling bertentangan secara diametris. Konsekuensinya, sangat penting untuk terlibat dalam asosiasi kalimat karena tidak cukup hanya mengandalkan kata-kata, misalnya dengan menggunakan parsing semantik.

Beberapa kata yang paling sering muncul diperoleh berdasarkan visualisasi yang telah dilakukan pada ulasan kelas sentimen positif dengan total 318 ulasan. Dalam hal ini, peneliti memunculkan 6 kata dengan frekuensi terbanyak, diantaranya kata "baik" yang digunakan 97 kali, kata "kirim" yang digunakan 85 kali, kata "cepat" yang digunakan 44 kali, kata "receive" yang digunakan sebanyak 37 kali, kata "courier" yang digunakan sebanyak 30 kali, dan kata "courier" yang digunakan sebanyak 30 kali.

Sementara itu, visualisasi sentimen negatif berdasarkan total 1.285 evaluasi mengungkapkan bahwa ada beberapa istilah yang paling sering muncul. Seperti dapat dilihat pada Gambar 4.17, enam kata teratas yang paling sering muncul telah direpresentasikan dalam contoh khusus ini. Kata "terkirim" muncul 614 kali, sedangkan istilah "paket" muncul 456 kali; kata "kecewa" muncul 345 kali; kata "buruk" muncul 201 kali; kata "parah" muncul 110 kali; dan kata "lambat" muncul 96 kali. Ini adalah kata-kata yang paling sering muncul dalam teks.

KESIMPULAN

Berikut adalah beberapa kesimpulan yang dapat ditarik berdasarkan pekerjaan yang telah dilakukan dari segi analisis dan pembahasan:

1. Ada 1.665 evaluasi pelanggan JNE yang diposting di situs web Google Play selama tanggal 2 Juni 2021 dan 30 September 2021. Jumlah pengguna terbanyak, 1.123, telah memberi item peringkat 1, yang sesuai dengan persentase sebesar 67,45%. Namun, setelah tahap preprocessing klasifikasi kelas emosi, ditemukan bahwa total 1.603 ulasan dapat diproses. Dari jumlah tersebut, total 1.285 ulasan negatif (80,16%) dan 318 ulasan positif (19,84%) dihasilkan.

2. Saat mempraktikkan metode *Naïve bayes Classifier*, lima cara berbeda untuk membandingkan data pelatihan dan data uji digunakan. Menurut temuan, rasio data pelatihan terhadap data uji adalah 90%:10%, yang menghasilkan tingkat akurasi, presisi, dan daya ingat tertinggi dengan akurasi 87,97%.

Informasi diperoleh bahwa, secara umum, pengguna mendiskusikan enam topik berbeda, termasuk pengiriman, paket, kurir, layanan, aplikasi, dan layanan. Informasi ini diperoleh berdasarkan hasil uji klasifikasi dan asosiasi. Pengguna memiliki banyak percakapan tentang kata "baik", "kirim", "cepat", "terima", "kurir", dan "layanan" di kelas yang berfokus pada emosi positif. Di sisi lain, orang yang mengekspresikan sikap negatif sering kali menggunakan kata dan frasa seperti "kirim", "paket", "kecewa", "mengerikan", "berat", dan "lambat".

Saya mengucapkan terima kasih kepada Ibu Neni sebagai pembimbing lapangan selama masa kerja praktik di PT. Tiki Jalur Nugraha Ekakurir Palembang. Terimakasih juga kepada bapak Irfan Dwi Jaya, M.Kom selaku dosen pembimbing lapangan di UIN Raden Fatah Palembang.

Referensi

- Choiriyah, Siti, & Setyo Riyanto. (2020). "Desain Pelatihan pada Masa Pandemi Covid19." *Syntax Idea 2.8* (2020): 488-502.
- Hariato & Saputra, A. (2020). Pengaruh Pengawasan Kerja Dan Disiplin Kerja Terhadap Kinerja Karyawan Pada Pt Centric Powerindo di Kota Batam. *Jurnal EMBA*. Vol.8 No.1 Februari 2020, Hal. 672-683
- Hasibuan, M. (2012). *Manajemen SDM*. Edisi Revisi, Cetakan Ke Tigabelas. Jakarta: Bumi Aksara.
- Hermanto. (2005). Analisis Pengaruh Pelatihan, Pendidikan dan Pembinaan Karyawan terhadap Kinerja Pegawai Sekretariat Kabupaten Sorong. *Jurnal Aplikasi Manajemen*, Vol. 3, No. 2, Hal. 131-139.
- Ivancevich, J. M., Robert, K. & Michael, T. M. (2007). *Perilaku dan Manajemen Organisasi*, Jakarta: Erlangga.
- Khurotin, N., & Afrianty, T. W. (2018). Analisis pelatihan dan Pengembangan Sumber Daya Manusia di PT Beon Intermedia Cabang Malang. *Jurnal Administrasi Bisnis*, 64(1), 195– 203.
- Kitchenham, B. and Charters, S. (2007) *Guidelines for Performing Systematic Literature Reviews in Software Engineering*, Technical Report EBSE 2007-001, Keele University and Durham University Joint Report.
- Lolowang, Melivin Grady dkk. (2016). Pengaruh Pelatihan dan Pengembangan Sumber Daya Manusia Terhadap Kinerja Karyawan Pada PT Berlian Kharisma Pasifik Manado. *Jurnal EMBA*, Vol. 4, No. 2, Hal. 177-186
- Mangkunegara, A. A Anwar prabu. (2017). *Manajemen Sumber Daya Manusia Perusahaan*, Bandung: Remaja Rosdakarya.
- Marjaya, I., & Fajar, P. (2019). "Pengaruh Kepemimpinan, Motivasi, Dan Pelatihan Terhadap Kinerja Pegawai." *Maneggio: Jurnal Ilmiah Magister Manajemen 2.1* (2019): 129-147.
- Choiriyah, Siti, & Setyo Riyanto. (2020). "Desain Pelatihan pada Masa Pandemi Covid19." *Syntax Idea 2.8* (2020): 488-502.
- Hariato & Saputra, A. (2020). Pengaruh Pengawasan Kerja Dan Disiplin Kerja Terhadap Kinerja Karyawan Pada Pt Centric Powerindo di Kota Batam. *Jurnal EMBA*. Vol.8 No.1 Februari 2020, Hal. 672-683

- Hasibuan, M. (2012). *Manajemen SDM*. Edisi Revisi, Cetakan Ke Tigabelas. Jakarta: Bumi Aksara.
- Hermanto. (2005). Analisis Pengaruh Pelatihan, Pendidikan dan Pembinaan Karyawan terhadap Kinerja Pegawai Sekretariat Kabupaten Sorong. *Jurnal Aplikasi Manajemen*, Vol. 3, No. 2, Hal. 131-139.
- Ivancevich, J, M., Robert, K. & Michael, T. M. (2007). *Perilaku dan Manajemen Organisasi*, Jakarta: Erlangga.
- Khurotin, N., & Afrianty, T. W. (2018). Analisis pelatihan dan Pengembangan Sumber Daya Manusia di PT Beon Intermedia Cabang Malang. *Jurnal Administrasi Bisnis*, 64(1), 195– 203.
- Kitchenham, B. and Charters, S. (2007) *Guidelines for Performing Systematic Literature Reviews in Software Engineering*, Technical Report EBSE 2007-001, Keele University and Durham University Joint Report.
- Lolowang, Melivin Grady dkk. (2016). Pengaruh Pelatihan dan Pengembangan Sumber Daya Manusia Terhadap Kinerja Karyawan Pada PT Berlian Kharisma Pasifik Manado. *Jurnal EMBA*, Vol. 4, No. 2, Hal. 177-186
- Mangkunegara, A. A Anwar prabu. (2017). *Manajemen Sumber Daya Manusia Perusahaan*, Bandung: Remaja Rosdakarya.
- Marjaya, I., & Fajar, P. (2019). "Pengaruh Kepemimpinan, Motivasi, Dan Pelatihan Terhadap Kinerja Pegawai." *Maneggio: Jurnal Ilmiah Magister Manajemen* 2.1 (2019): 129-147.