



Shopee Application Review Classification Using Probabilistic Neural Network Algorithm And K-Nearest Neighbor

Klasifikasi Ulasan Aplikasi Shopee Menggunakan Algoritma Probabilistic Neural Network Dan K-Nearest Neighbor

M. Afdal¹, Lia Waroka²

^{1,2}Department of Information System, Faculty of Science and Technology
UIN Sultan Syarif Kasim Riau

Email: m.afdal@uin-suska.ac.id liawaroka@gmail.com

Makalah: Diterima 21 Januari 2022; Diperbaiki 28 Maret 2022; Disetujui 28 Maret 2022
Corresponding Author: Name: M. Afdal

Abstrak

Google Play Store merupakan layanan konten digital yang dimiliki oleh Google yang melingkupi toko *online*. Peningkatan unduhan aplikasi shopee pada *google play store* membawa berbagai macam komentar yang di berikasikan oleh semua orang yang menggunakan aplikasi shopee meliputi komentar yang positif, netral atau negatif mengenai aplikasi shopee tersebut. Penelitian ini dilakukan bertujuan untuk melihat perbandingan menggunakan 2 algoritma klasifikasi yaitu algoritma Probabilistic Neural Network (PNN) dan K-Nearest Neighbor (KNN). Pembagian data penelitian ini menggunakan K-Fold Cross Validation yang kemudian dihitung akurasi pada data ulasan aplikasi dan produk Shopee. Akurasi data ulasan aplikasi KNN lebih tinggi dibandingkan PNN dengan akurasi KNN sebesar 77,85% sedangkan PNN dengan akurasi 72,43%. Untuk data produk akurasi KNN lebih tinggi dibanding PNN dengan akurasi KNN sebesar 91,43% sedangkan PNN sebesar 85,71%. Jadi hasil akurasi pada data yang digunakan dalam penelitian ini, algoritma KNN memiliki performa lebih baik dibanding PNN.

Kata Kunci : Google Play Store, K-Nearest Neighbor, Probabilistic Neural Network, Shopee, Text Mining.

Abstract

Google Play Store is a digital content service owned by Google that covers online stores. The increase in downloads of the shopee application on the google play store brings a variety of comments that are commented on by everyone who uses the shopee application including optimistic, neutral, or pessimistic comments on the shopee application. This study was conducted to see a comparison using 2 classification algorithms, namely the Probabilistic Neural Network (PNN) and K-Nearest Neighbor (KNN) algorithms. The data sharing of this study uses K-Fold Cross Validation which is then calculated for accuracy on Shopee application and product review data. The accuracy of KNN application review data is higher than PNN with a KNN accuracy of 77.85% while PNN with an accuracy of 72.43%. For product data, the accuracy of KNN is higher than PNN with a KNN accuracy of 91.43% while PNN is 85.71%. So the results of accuracy in the data used in this study, the KNN algorithm has a better performance than PNN.

Keyword: Google Play Store, K-Nearest Neighbor, Probabilistic Neural Network, Shopee, Text Mining.

1. Pendahuluan

Saat ini belanja *online* sudah menjadi kegiatan yang dipilih oleh masyarakat agar tidak keluar rumah dan praktis untuk berbelanja. Hal ini menjadi salah satu faktor terjadinya perubahan konsumen ke arah digital dan memanfaatkan layanan *online* yang disediakan seperti *e-commerce* dan media sosial [1]. Saat ini toko *online* (*marketplace*) Shopee juga menjadi layanan *e-commerce* yang paling sering digunakan dan urutan nomor 1 di Indonesia, menurut survei yang dilakukan *DailySocial* untuk mengetahui layanan *e-commerce* favorit atau yang paling banyak digunakan versi responden survei yang dilakukan terhadap 2.026 responden diseluruh Indonesia yang menggunakan *marketplace*, menyatakan bahwa Shopee menjadi layanan *e-commerce* yang paling sering digunakan oleh responden yaitu sebanyak (33,63%), kemudian posisi berikutnya berturut-turut diikuti Tokopedia dengan angka persentase (28,11%), Bukalapak (17,50%), Lazada (14,30%), Blibli (2,15%), serta pilihan *e-commerce* yang lainnya (4,31%) [2].

Salah satu situs belanja *online* yang paling diminati oleh masyarakat Indonesia yakni Shopee. Shopee sendiri merupakan situs belanja *online* yang paling diminati dan merupakan *E-commerce* terbesar di Asia Tenggara[3]. Shopee telah berkembang cukup pesat dari *marketplace* lainnya. Saat ini shopee telah di unduh sebanyak 100 juta lebih dan memiliki 8 juta ulasan pada *Google Play Store*. Setiap harinya shopee telah melayani lebih dari 200.000 transaksi. pada tahun 2017 shopee juga sempat menjadi *top number one shopping app* di *Google Play Store*[4].

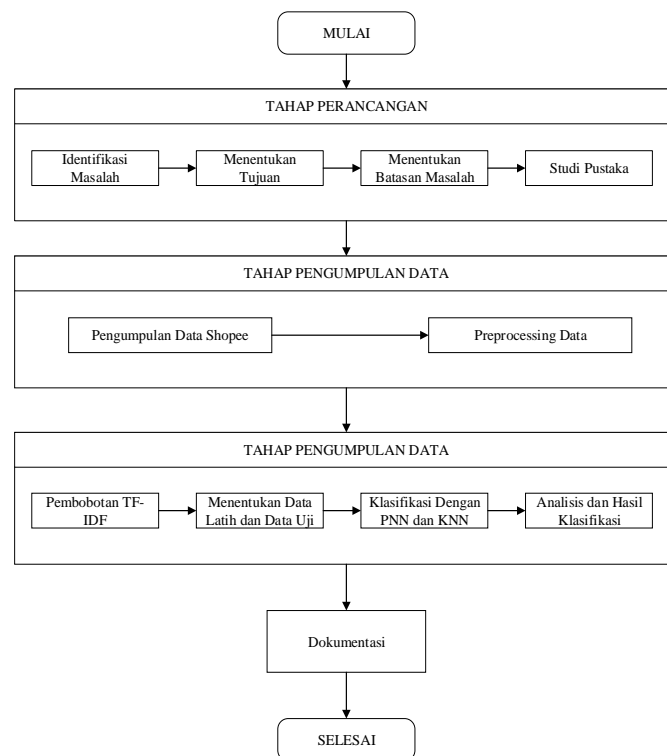
Tahun 2017 shopee juga sempat menjadi *top number one shopping app* di google play store (Angraini, Perbawasari, and Budiana 2018). Peningkatan unduhan app shopee pada google play store bagi pengguna menuai banyaknya komentar-komentar atau ulasan mengenai aplikasi Shopee pada ulasan di google play store membawa berbagai macam pengaruh yang besar bagi pihak Shopee dan pengguna karena meliputi komentar yang positif, netral atau negatif mengenai aplikasi shopee tersebut. Salah satu cara untuk mengevaluasi komentar-komentar pengguna pada aplikasi ini dengan menggunakan analisa sentimen. Metode yang akan digunakan dalam analisa sentimen penelitian ini yaitu *probabilistic neural network* dan *k-nearest neighbors*.

2. Metodologi

Algoritma *k-nearest neighbor* (k-nn) merupakan salah satu algoritma yang paling populer digunakan pada machine learning hal ini karena prosesnya mudah dan sederhana[5]. Metode Knn data kelompok baru berdasarkan dari estimasi jarak data baru ke beberapa tetangga terdekat lainnya[6]. Algoritma *probabilistic neural network* (pnn) merupakan metode yang sering digunakan penelitian dalam *dataset* gambar maupun teks dan menghasilkan akurasi baik seperti pada kasus analisis klasifikasi[7].

Salah satu model pada Jaringan Syaraf Tiruan (JST) yang digunakan untuk pengklasifikasian *Probabilistic neural network* (pnn) merupakan suatu metode (JST) yang menggunakan prinsip dari teori statistik yaitu *Bayesian Classification* untuk menggantikan prinsip heuristik yang digunakan oleh algoritma *back propagation*[8].

Langkah-langkah yang akan ditempuh dalam Penelitian ini dapat di lihat pada Gambar 1 metode penelitian sebagai berikut:



Gambar 1. Metodologi Penelitian

2.1 Tahap Perencanaan

Berikut merupakan tahap perencanaan dari metode penelitian :

1. Identifikasi Masalah

Mengidentifikasi masalah yaitu dari mengamati permasalahan terdahulu apa yang terjadi berdasarkan komentar aplikasi shopee pada *google play store* tersebut. Dengan menggunakan algoritma *probabilistic neural network (pnn)* dan *k-nearest neighbor (knn)*.

2. Penentuan tujuan

Menentukan tujuan untuk memperjelas maksud yang harus dilakukan. Tujuan dari penelitian ini adalah untuk mengklasifikasikan *sentiment public* atau komentar masyarakat terhadap aplikasi shopee yang mereka gunakan, dari *google play store*, kemudian menghitung akurasi algoritma yang digunakan pada data tersebut yaitu *probabilistic neural network* dan *k-nearest neighbor*.

3. Penentuan Batasan Masalah

Penetapan batasan masalah penelitian bertujuan agar penelitian tidak berada di luar lingkup objek penelitian yang telah ditentukan.

4. Studi Pustaka

Studi pustaka bertujuan untuk mengetahui teori-teori yang berkaitan dengan penelitian, kemudian digunakan untuk memecahkan masalah yang ada dalam penelitian. Studi pustaka juga berguna sebagai dasar referensi yang kuat bagi peneliti untuk menganalisis opini masyarakat tentang ulasan aplikasi shopee di *google play store* dengan menggunakan algoritma *probabilistic neural network* dan *k-nearest neighbor*.

2.2 Tahap Pengumpulan Data

1. Pengumpulan data ulasan aplikasi shopee

Pada pengumpulan data ulasan aplikasi shopee ini, akan dilakukan *crawling* data dengan menggunakan *google collab* yang berasal dari ulasan aplikasi shopee pada *google play store*.

2. Pelabelan Data

Pelabelan data yaitu menentukan kumpulan kata tersebut apakah bermakna negatif, netral, atau positif dengan melihat dari pemilihan kata dan makna yang disampaikan pada opini tersebut secara manual.

3. Preprocessing

Berikut ini tahap *Text Preprocessing* yang digunakan:

A. *Cleaning*

Langkah *cleaning* untuk menghapus karakter selain huruf, mengubah teks menjadi huruf kecil (*case folding*), menghapus *username* pengguna atau mentions, menghapus hashtag (#), menghapus URL atau link dari setiap ulasan aplikasi shopee dan menghapus yang sama.

B. *Filtering*

Tahap *filtering* tahap ini dimana setiap mengambil kata-kata penting yang telah 21 dipotong kata per kata dari sebuah kalimat. Pada tahap ini dilakukan pembuangan kata-kata yang dianggap tidak penting seperti tanda baca dan lainnya.

C. *Tokenizing*

Tahap *tokenizing* tahap ini pemotongan input string berdasarkan setiap kata yang menyusunnya, disini teks yang dikirim dipotong untuk dipisahkan untuk diproses lebih lanjut.

D. *Stemming*

Tahap *stemming* tahap ini dimana untuk menemukan akar kata dari setiap hasil filtering. Dokumen atau artikel terdiri dari rangkaian kalimat yang saling terkait. Dalam setiap kalimat terlihat jelas bahwa pola kalimat dasar paling tidak berdasarkan SPO (*Subjek, Predikat, Objek*).

3. Hasil dan Analisis

Analisa kebutuhan data merupakan proses identifikasi serta pembahasan hal yang dilakukan sebelum membuat penyelesaian yang nantinya akan menjadi hasil utama dari penelitian ini. Serta proses identifikasi penelitian serta membahas tahapan-tahapan apa saja yang akan dilakukan yang nantinya akan menunjukkan hasil dari penyelesaian penelitian ini. Data yang digunakan pada penelitian ini yaitu data dari ulasan pengguna aplikasi

shopee di *google play store* dengan kata kunci yang di ambil “komentar aplikasi dan produk”. Dengan rentang waktu mei 2021 sampai dengan oktober 2021.

3.1 Pengumpulan Data

Tahapan pengumpulan data dilakukan untuk mendapatkan data yang akan digunakan dalam penelitian. Pengumpulan data pada penelitian ini dilakukan dengan cara crawling pada ulasan aplikasi shopee di *Google Play Store* dengan menggunakan *Google Collab* terdapat tabel 1.

Table 1. Data Awal Ulasan Aplikasi Shopee

No	Nama	Tanggal	Ulasan aplikasi dan produk shopee
1	lyly indrayani	17/06/2021	Untuk pelayanan order nya mah kerencepet sampe barang nya, real pict.. Cuma saya mengajukan SPAYLATER sudah 3kali di tolak terus, tolong dong kalo EMGgak di ACC jangan ngasih ajukan SPAY- LATER nya.. Bikin ngasih harapan pal-su aja, padahal saya sudah masuk member gold.
2	nana ratna	17/06/2021	Saya mau jual barang tapi kenapa kogbisa buat toko.. saat mau jual suruh pilihjasa pengiriman tapi setelah dipilih harusmasukin alamat, udah di masukin alamattpi gak bisa,, malah suruh masukin alamatlagi, lagi dan lagi dan gak bisa juga.
3	m arda bili	17/06/2021	adain bintang nol kalo bisa, malesngasih bintang pengemasan barang dishoppe s-mart sampe 10hari gila kali,orang beli barang nungu selama itucuma buat pengemasan 10hari penge-masan dikit doang udah kaya ngemas pe-sawat,,buang2 waktu banget belanja dis-itu, ditunggu sampe 10hari pengemasan,udah kaya beli barang dari luar negeri,tapi masih mending,ini mah ga ada perge-rakan sama sekali, pengemasan sama pe-ngiriman lemot, bikin kecewa pelanggan,Payah.
4	sopy bidin	17/06/2021	Saya beli hp, tapi baru seminggu hpmati total, chat ke penjual, penjual malahmenghindar, minta bantuan shopee, gabisa karena udah klik pesan di terima,padahal di deskripsi ada garansi 1 tahun,buat pelajaran aja, kalau beli baju ataumainan, bolehlah di toko online, tapi un-tuk barang elektronik, bakal kecewa, baruseminggu rusak, ga bisa apa2, walau ter-tera garansi 2 tahun, tapi saat minta garan-si, ga di tanggapi.
5	mamiwa	17/06/2021	Demi kenyamanan? Isi shopeepay sekarang ada biaya admin 500. Blm potongan admin untuk tiap pesanan. Untung seller ga seberapa. Naikin harga, jarang yg beli. Bahan baku apa lagi~~~ lagi susah. Gratis ongkir juga ada syaratnya harus belanja segini segitu... Shopee~~~ ga mau cekik aku aja nihhh????
6	gina paripurna	17/06/2021	Bertahun tahun pake aplikasi shopee, sekarang kecewa banget!, produk luar negri udah dihapus sama shopee pdahal seller lokal yg jualan online kbanyakan ambil produk impor, produk di keranjang belanja yang impor juga dihapus.. di marketplace lain boleh boleh aja tuh produk luar.. jgn sampe nanti pengguna shopee pindah ke

apk lain ya tolong di kembalikan spt semula lagi shopee
mohon maaf

...
3020	suwarno	10/06/2021	Shopee adalah sebuah aplikasi yang sangat bagus, berbelanja mudah serta fitur-fiturnya yang beragam dan menarik..

Setelah dilakukan *Cleaning* dari 3020 data yang didapat hanya 510 data yang bisa digunakan, terlihat seperti Tabel2.

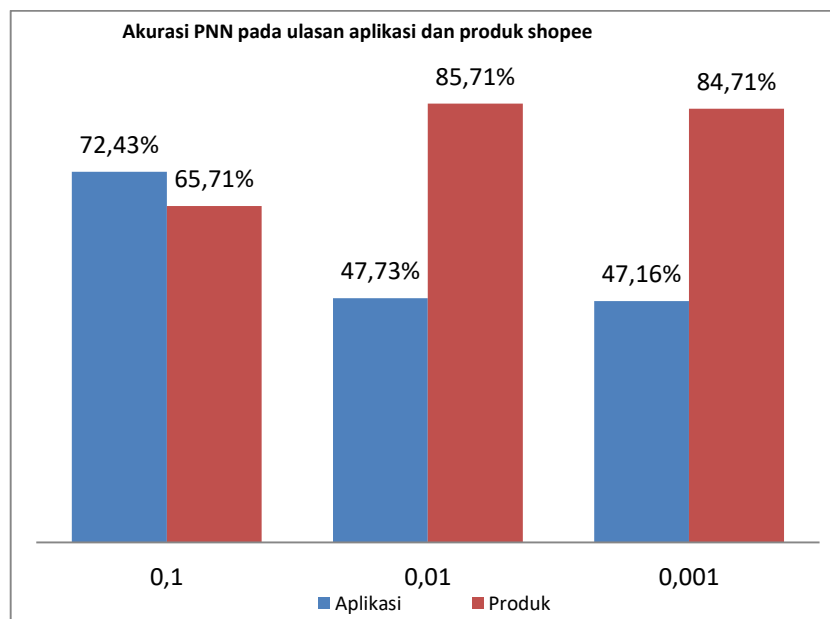
Table 2. Hasil *cleaning* data ulasan produk shopee

No	Ulasan Produk
1	masya allah shoppe terbaik memang pesan apa apa selalu ada dengan harga yang terjangkau biasanya belanja online selalu takut ga sesuai tp disini selalu memuaskan semoga shoppe terus maju menjadi pusat perbelanjaan no di indonesia
2	aku beli gorden di shopee yang datang cuma kardus bekas pas dilaporin ko pihak shopee malah ga ada respon rugi an nasib belanja online ya gini skarang aku ga pake shopee lagi males mending pake e commerce yg lain biarpun bayar ongkir mahal yang penting aman
3	bonus jt dari hongkong gak niat berbagi hadiah gak usah sok sok an kasih hadiah saya undang temen saya yang bari beli hp suruh pake link saya keterangannya ada yang janggal lucu nih aplikasi sampah gak usah di download buang waktu dan kuota
4	untuk pelayanan order nya mah keren cepet sampe barang nya real pict cuma saya mengajukan spaylater sudah kali di tolak terus tolong dong kalo emg gak di acc jangan ngasih ajukan spaylater nya bikin ngasih harapan palsu aja padahal saya sudah masuk member <i>gold</i>
5	saya mau jual barang tapi kenapa kog bisa buat toko saat mau jual suruh pilih jasa pengiriman tapi setelah dipilih harus masukan alamat udah di masukan alamat tpi gak bisa malah suruh masukan alamat lagi lagi dan lagi dan gak bisa juga
6	saya beli hp tapi baru seminggu hp mati total chat ke penjual penjual malah menghindar minta bantuan shopee ga bisa karena udah klik pesan di terima padahal di deskripsi ada garansi tahun buat pelajaran aja kalau beli baju atau mainan bolehlah di toko online tapi untuk barang elektronik bakal kecewa baru seminggu rusak ga bisa apa walau tertera garansi tahun tapi saat minta garansi ga di tanggapi
7	demi kenyamanan isi shopeepay sekarang ada biaya admin blm potongan admin untuk tiap pesanan untung seller ga seberapa naikin harga jarang yg beli bahan baku apa lagi lagi susah gratis ongkir juga ada syaratnya harus belanja segini segitu shopee ga mau cekik aku aja niiihhh

- 8 kenapa barang luar negeri di batasi kenapa ga dari awal aja shopee indonesia barang lokal semua g diperbolehkan barang luar klo gt buat apa ada filter luar negeri kalo tiap liat barang luar terus ada text barang dibatasi padahal rezeki masing lagian ada alasan tersendiri yang suka belanja brg luar barang luar suka real pict terus murah termasuk ongkir kalo barang lokal jarang real pict terus harga nya suka kli lipat dan bnyk toko yg cheat g mau kmbaliin dana tlg jgn dibatasi lgi yee
- 9 sdh mulai ragu sya pesan beberapa kali pesan sepatu sizenya salah dibilangnya katanya redy pesan size yg datang kalau barangnya kosong jgn dipaksakan ada tpi ukuran salah bgini kan bikin kecewa pelanggan
- 10 tolong pihak shopee supaya bisa membuka blokir produk yang dari luar negeri yang memang produk nya tidak ada di indonesia karena tidak semua nya indonesia bisa menyediakan barang nya saya sangat kecewa dengan kebijakan ini
- ...
- 510 ga suka pilihan produk nya ga banyak

3.2 Proses Metode PNN

Probabilistic Neural Network (PNN) merupakan suatu metode jaringan syaraf tiruan yang menggunakan prinsip dari teori statistik yaitu bayesian classification untuk menggantikan prinsip heuristik yang digunakan oleh algoritma *back propagation* dapat dilihat pada gambar 2.



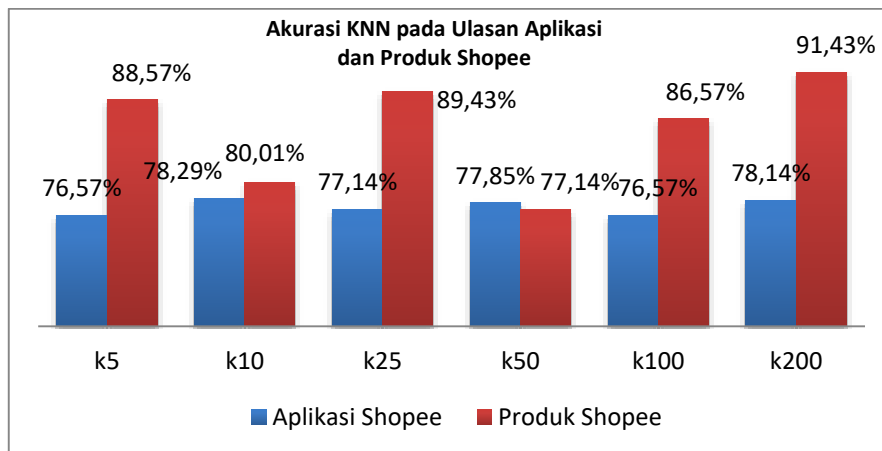
Gambar 2. Akurasi PNN Antara Data Ulasan Tentang Aplikasi Dan Produk Shopee

Perbandingan akurasi PNN antara data ulasan tentang aplikasi dan produk shopee yang tertinggi adalah data Produk dengan akurasi sebesar 85.71% pada nilai spread = 0.01, sedangkan data Aplikasi sebesar 72,43% pada spread = 0.1. Hasil perbandingan akurasi PNN dengan menggunakan pembagian data *K-Fold Cross Validation* pada data tentang ulasan aplikasi dan produk shopee.

3.3 Proses Metode KNN

K-Nearest Neighbor (KNN) merupakan algoritma yang sering digunakan dalam klasifikasi data, algoritma ini juga digunakan untuk mengklasifikasikan data yang terdekat dari objek tersebut. Penerapan algoritma KNN pada penelitian ini digunakan untuk mendapatkan akurasi klasifikasi *text mining* dengan menggunakan data

ulasan dari pengguna aplikasi dan produk shopee pada google play store. KNN dalam penelitian ini menggunakan rumus jarak *Euclidean Distance* dapat dilihat pada gambar 3.

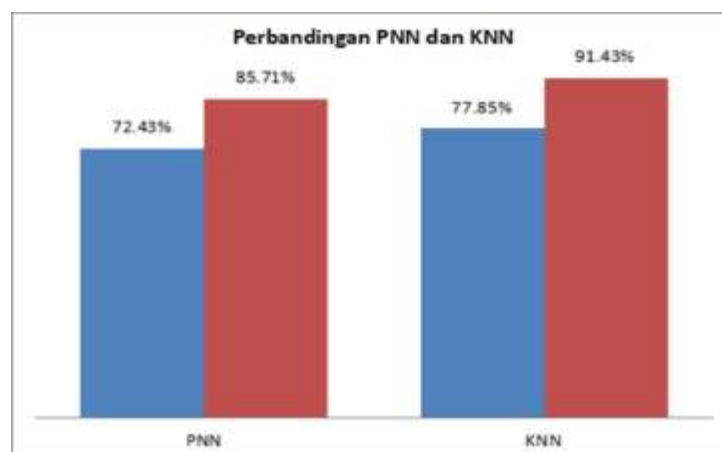


Gambar 3. Akurasi KNN Antara Data Ulasan Tentang Aplikasi Dan Produk

KNN dalam penelitian ini menggunakan rumus jarak *Euclidean Distance*. Perbandingan akurasi KNN antara data ulasan aplikasi dan produk shopee yang tertinggi adalah data ulasan mengenai produk shopee dengan akurasi sebesar 91,43% pada parameter $k = 200$, sedangkan data ulasan aplikasi shopee sebesar 78,85% pada parameter $k = 50$. Hasil perbandingan akurasi KNN dengan menggunakan pembagian data *K-Fold Cross Validation* pada ulasan aplikasi dan produk shopee.

3.4 Perbandingan Hasil Metode PPN dan KNN

Dari percobaan yang telah dilakukan di dapatkan 2 rekomendasi teratas pada percobaan dengan menggunakan metode PNN dan KNN yang dapat dilihat pada gambar 4.



Gambar 4. Perbandingan Hasil Perengkingan

3.5 Visualisasi Data

Visualisasi kata dengan memanfaatkan salah satu situs *webwordclouds* yaitu <https://wordart.com/>. Hasil dari visualisasi tersebut dapat melihat jumlah frekuensi kata pada opini netral, negatif dan positif berdasarkan data ulasan aplikasi shopee pada *google play store* tentang aplikasi dan produk dari yang telah didapat sesuai dengan kelasnya masing-masing.



Gambar 5. Visualisasi Opini Positif Data Ulasan Aplikasi Shopee

Berdasarkan Gambar 5 dapat dilihat bahwa kata yang sering muncul pada opini positif data Aplikasi yaitu kata “Shopee” dan disusul oleh kata lain.



Gambar 6. Visualisasi Opini Negatif Data Ulasan Produk Shopee

Berdasarkan Gambar 6 dapat dilihat bahwa kata yang sering muncul pada opini negatif data Produk yaitu kata “Saya” dan disusul oleh kata lainnya.



Gambar 7. Visualisasi Opini Netral Data Ulasan Produk Shopee

Berdasarkan Gambar 7. dapat dilihat bahwa kata yang sering muncul pada opini netral data Produk yaitu kata “Di” dan disusul oleh kata lainnya.

4. Kesimpulan

Pada penelitian ini metode pembagian data menggunakan *K-Fold Cross Validation* didapat hasil perbandingan akurasi antara PNN dan KNN terhadap data ulasan aplikasi Shopee pada Google Play Store tentang aplikasi dan produk terhadap pengguna Shopee, Dimana pada data aplikasi akurasi KNN lebih tinggi dibandingkan PNN dengan akurasi KNN sebesar 77.85% sedangkan PNN sebesar 72.43%. Untuk data Produk akurasi KNN juga lebih tinggi dibandingkan PNN dengan akurasi KNN sebesar 91,43% sedangkan PNN sebesar 85.71%. Jadi hasil akurasi aplikasi Shopee pada *Google Play Store* tentang ulasan aplikasi dan produk oleh pengguna menggunakan algoritma KNN memiliki performa lebih baik dari PNN.

5. Ucapan Terimakasih

Ucapan terima kasih untuk keluarga besar Prodi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah menjadi wadah dalam menggali ilmu pengetahuan dan senantiasa memberikan dukungan dan motivasi hingga penelitian ini dapat terselesaikan sesuai dengan harapan.

Referensi

- [1] Amalia, H., Retnasari, T., dan Rachmawati, S. (2020). Pemanfaatan teknologi informasi untuk meningkatkan pelayanan akademik rumah tahfidz dan tpq sakinah cipayung jakarta timur. *Jurnal Abdimas BSI: Jurnal Pengabdian Kepada Masyarakat*, 3(2), 228–235.
- [2] Hanifah, N., dan Rahadi, D. R. (2020). Analisis perilaku konsumen dalam memusulkan pembelian secara online pada masa pandemi covid-19. *SULTANIST: Jurnal Manajemen dan Keuangan*, 8(2), 112–122.
- [3] Agianto, R., Setiawati, A., dan Firmansyah, R. (2020). Pengaruh media sosial instagram terhadap gaya hidup dan etika remaja. *TEMATIK Jurnal Teknologi Informasi Dan Komunikasi*, 7(2), 130–139.
- [4] Anggraini, C. A., Perbawasari, S., dan Budiana, H. R. (2018). Cyberbranding sebagai upaya membangun brand awareness shopee indonesia: Chindy ayu angraini, susie perbawasari, heru ryanto budiana. *Commed: Jurnal Komunikasi dan Media*, 2(2), 72–86.
- [5] Firman Tempola, M. M., dan Khairan, A. (2018). Perbandingan klasifikasi antara knn dan naive bayes pada penentuan status gunung berapi dengan k-fold cross validation. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 5(5).
- [6] Wiyanto, W., Ngudi, T., dan Saefulloh, A. (2020). Analisa tingkat kepuasan pelanggan terhadap pelayanan perusahaan otobus xyz menggunakan metode naive bayes. *Pelita Teknologi*, 15(1), 56–67.
- [7] Najiyah, I., dan Hariyanti, I. (2021). Sentimen analisis covid-19 dengan metode probabilistic neural network dan tf-idf. *Jurnal Responsif: Riset Sains & Informatika*, 3(1), 100–111.
- [8] Adyati, R. D., Nasution, Y. N., dan Wahyuningsih, S. (2019). Klasifikasi probabilistic neural network (pnn) pada data diagnosa penyakit demam berdarah dengue (dbd) tahun 2018. Dalam *Prosiding seminar*

nasional matematika dan statistika (hal. 15–21).