

Klasifikasi Sentimen untuk Analisis Kepuasan Pelayanan Puskesmas Berbasis Arsitektur LSTM

I Gede Bintang Arya Budaya^{*1}, Luh Putu Safitri Pratiwi², Dedy Panji Agustino³

¹Program Studi Teknologi Informasi, Institut Teknologi dan Bisnis STIKOM Bali

^{2,3}Program Studi Sistem Informasi, Institut Teknologi dan Bisnis STIKOM Bali

E-mail: ^{*1}bintang@stikom-bali.ac.id, ²putu_safitri@stikom-bali.ac.id,

³panji@stikom-bali.ac.id

Abstrak

Transformasi layanan sistem kesehatan di Indonesia menjadi fokus utama pemerintah melalui program nasional satu data kesehatan Indonesia. Dalam program ini, digitalisasi teknologi kesehatan menjadi salah satu pilar transformasi yang penting. Puskesmas, sebagai pusat layanan kesehatan level pertama, menjadi salah satu fokus dalam transformasi ini. Kepuasan pelayanan di Puskesmas menjadi faktor penting dalam mencapai tujuan tersebut. Penelitian ini bertujuan untuk menganalisis pemanfaatan teknik klasifikasi sentimen dalam survei kepuasan pelayanan Puskesmas. Dalam penelitian ini, diterapkan pendekatan deep learning dengan menggunakan arsitektur Long-Short Term Memory (LSTM) untuk klasifikasi sentimen. Hasil penelitian menunjukkan bahwa model LSTM efektif digunakan untuk mengklasifikasikan sentimen. Akurasi keseluruhan model mencapai 76.26%, menunjukkan kemampuan model dalam melakukan prediksi dengan baik. Namun, terdapat perbedaan performa antara kelas sentimen. Sentimen netral memiliki performa yang sedikit lebih rendah dibandingkan dengan sentimen positif dan negatif. Hal ini dapat disebabkan oleh ketidakseimbangan jumlah sampel, kompleksitas sentimen netral, variabilitas dalam teks, dan perbedaan subyektivitas dalam penilaian sentimen. Penelitian ini menunjukkan potensi penerapan deep learning, khususnya model LSTM, dalam analisis sentimen kepuasan pelayanan Puskesmas. Hasil penelitian ini dapat menjadi dasar untuk integrasi deep learning dalam pengembangan sistem digital kesehatan nasional.

Kata Kunci—prediksi sentiment, pelayanan puskesmas, deep learning.

1. PENDAHULUAN

Dalam upaya mentransformasi layanan sistem kesehatan di Indonesia, pemerintah melalui kementerian kesehatan republik Indonesia menginisiasi program nasional satu data kesehatan Indonesia. Dalam program nasional ini terdapat enam pilar transformasi yang menjadi arah dalam perwujudan program, salah satunya adalah pilar transformasi melalui digitalisasi teknologi kesehatan. Pusat kesehatan masyarakat (Puskesmas) merupakan salah satu entitas yang menjadi fokus dalam transformasi layanan kesehatan ini. Puskesmas sebagai pusat kesehatan level pertama yang fungsinya paling dekat dengan masyarakat, kepuasan dalam pelayanan kesehatan tentu menjadi poin penting yang pada akhirnya dapat tercapai jika kebutuhan pelayanan standar pasien dapat terpenuhi [1]. Oleh karena itu secara berkala dilaksanakan survei kepuasan terhadap pelayanan puskesmas oleh pihak – pihak terkait.

Pengembangan *big data* menjadi dasar utama untuk dapat mewujudkan sistem satu data kesehatan Indonesia [2]–[4], dan tentu survei kepuasan terhadap puskesmas harusnya juga terintegrasi di dalam sistem kesehatan tersebut. Wawasan yang didapatkan dari hasil analisa survei kepuasan selanjutnya dapat dimanfaatkan untuk pengambilan keputusan dalam upaya peningkatan kualitas pelayanan di puskesmas tersebut. Penelitian ini bertujuan untuk menganalisa pemanfaatan teknik klasifikasi untuk mengetahui sentiment dari hasil survei kepuasan yang

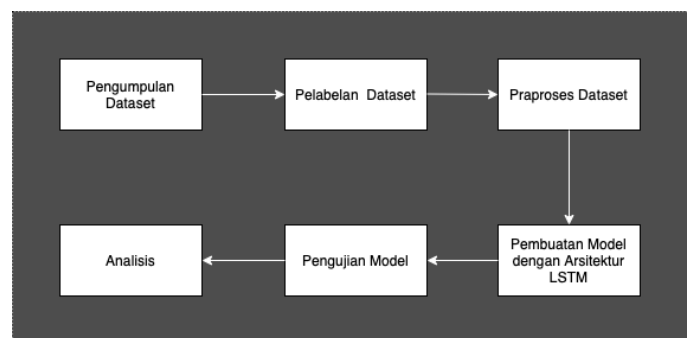
didapatkan. Data survei yang digunakan bersumber dari laporan survei kepuasan pelayanan puskesmas di kota Denpasar yang dilaksanakan oleh *Center of Public Health Innovation (CPHI)* dari universitas Udayana, Denpasar, Bali.

Beberapa penelitian telah dilaksanakan dalam hal klasifikasi sentimen, seperti analisa sentimen yang bertujuan untuk memahami persepsi publik terkait pelayanan dari suatu lembaga berdasarkan persepsi dari pengguna media sosial [5]. Penelitian lainnya menganalisa berdasarkan hasil penilaian yang diberikan pengguna secara langsung terhadap suatu produk atau jasa yang dimiliki oleh suatu lembaga [6]. Proses analisis sentimen dapat dilakukan secara sederhana dengan membagi sentimen tersebut menjadi dua kelas, positif dan negatif, serta ada juga yang membaginya menjadi lebih dari dua kelas dengan menggunakan *machine learning* [7] maupun *deep learning* [6]. Dalam hal analisis sentimen untuk Puskesmas, beberapa penelitian juga telah dilaksanakan [7], [8], dalam kasus penelitian tersebut analisis sentimen memiliki lebih dari dua kelas dan berfokus pada eksplorasi performa dari pemanfaatan algoritma *Support Vector Machine (SVM)*.

Penelitian ini akan berfokus pada potensi penerapan *deep learning* dalam kasus kepuasan pelayanan Puskesmas dengan memanfaatkan arsitektur *Long-Short Term Memory (LSTM)* [6], [9], [10]. LSTM digunakan untuk menjalankan proses klasifikasi sentimen yang didapatkan dari hasil survei kepuasan pelanggan dari hasil pelayanan Puskesmas khususnya di kota Denpasar. Hasil klasifikasi sebagai bentuk gambaran umum sentimen yang diterima oleh pelayanan kesehatan Puskesmas terkait (positif, negatif, dan netral) yang selanjutnya dapat dianalisa lebih dalam mengenai relevansi pelayanan yang menyebabkan sentimen masuk dalam klasifikasi tersebut. Harapannya hasil dari penelitian ini dapat menjadi salah satu fondasi serta pembandingan dalam potensi penerapan dan integrasi *deep learning* dalam pengembangan sistem digital kesehatan nasional.

2. METODE PENELITIAN

Penelitian dilaksanakan melalui beberapa tahapan. Adapun metode dalam pelaksanaan penelitian ini dapat terlihat seperti yang ada pada Gambar 1 berikut.



Gambar 1. Metode Penelitian

2.1. Pengumpulan Dataset

Dataset berasal dari laporan survei kepuasan pelayanan Puskesmas di kota Denpasar dan Kabupaten Gianyar yang dilaksanakan oleh CPHI fakultas kedokteran universitas Udayana, Denpasar, Bali [11]. Dalam dataset ini terdiri dari opini yang diberikan oleh pasien pasca mendapatkan pelayanan dari Puskesmas terkait. Opini tidak hanya sebatas pelayanan kesehatan, namun juga termasuk pelayanan pendukung dalam proses pemanfaatan fasilitas kesehatan di

Puskesmas terkait. Terdapat 703 baris opini berdasarkan laporan survei tersebut. Adapun sampel dari opini tersebut dapat dilihat pada Tabel 1 berikut.

Tabel 1. Sampel Dataset Hasil Survei Kepuasan Pelanggan

No	Opini
1	Ruang tunggu kursi mohon ditambah
2	Sudah bagusmohon ditingkatkan
...	...
701	Waktu tunggu sangat lama

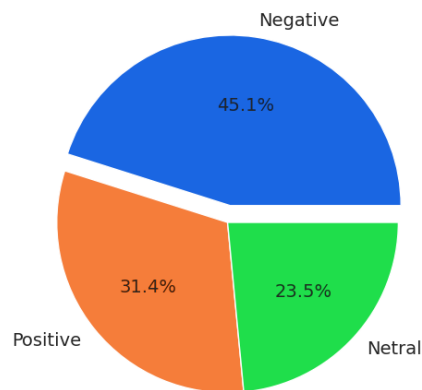
2.2. Pelabelan Dataset

Pada proses pelabelan dataset dilakukan oleh ahli yang terdiri dari dua orang ahli untuk memberikan label sentimen untuk setiap opini yaitu positif, negatif, atau netral. Setiap ahli tersebut memberikan label masing – masing pada keseluruhan dataset, dimana setelah pemberian label, hasil pelabelan dijadikan satu. Jika dalam proses pelabelan antara kedua ahli memiliki nilai label yang berbeda, maka opini tersebut akan dihapus, hanya opini yang memiliki label sama oleh kedua ahli tetap digunakan. Hal tersebut dilakukan untuk meminimalisir bias, dimana dari 701 opini tersisa 694 opini yang digunakan. Adapun sampel opini dengan labelnya dapat dilihat pada Tabel 2 berikut serta rasio persentase setiap kelas label ada pada

Gambar 2 berikut.

Tabel 2. Sampel Dataset Hasil Pelabelan

No	Opini	Sentimen
1	Ruang tunggu kursi mohon ditambah	Netral
2	Sudah bagusmohon ditingkatkan	Positif
...	...	...
694	Waktu tunggu sangat lama	Negatif



Gambar 2. Rasio Setiap Label Kelas (%)

2.3. Praproses Dataset

Pada tahap praproses dataset, dataset melalui beberapa proses, pertama adalah memperbaiki penulisan yang salah sebagai contoh penulisan “*bagusmohon*” pada sampel dataset seharusnya menjadi “*bagus mohon*”, terdapat penambahan spasi sehingga data yang dibaca sebagai satu kata sebelumnya menjadi 2 kata yang memang tepat. Selanjutnya adalah praproses dataset standar seperti *case folding*, penghilangan tanda baca, penghilangan angka, tokenisasi dan

stemming [12], [13] . Proses *stemming* untuk merubah kata menjadi kata dasar menggunakan *library* Sastrawi [14], [15]. Dataset final selanjutnya dibagi menjadi data latih sebanyak 80% dan data uji sebanyak 20%. Tabel 3 menampilkan hasil praproses data.

Tabel 3. Sampel Hasil Praproses Dataset

No	Opini	Hasil Praproses
1	Ruang tunggu kursi mohon ditambah	[ruang, tunggu, mohon, tambah]
2	Sudah bagusmohon ditingkatkan	[sudah, bagus, mohon, tingkat]
...	...	...
694	Waktu tunggu sangat lama	[waktu, tunggu, sangat, lama]

2.4. Pembuatan Model LSTM

Model LSTM telah terbukti efektif dalam analisis sentimen dengan kemampuannya dalam memahami konteks dan ketergantungan dalam teks. Penggunaan model LSTM dalam analisis sentimen telah membantu dalam menganalisis dan mengklasifikasikan sentimen dalam teks dengan cukup akurat [9], [10]. Dalam implementasi arsitektur yang digunakan pada penelitian ini menggunakan *sequential model* dari *library Keras*. Model memiliki parameter latih seperti *batch size* = 128, *optimizer Adam*, fungsi *loss* dengan *sparse categorical cross entropy* dan jumlah *epoch* untuk melatih model adalah 1000.

2.5. Pengujian Model

Model diuji dan dibuat *confusion matrix* serta dilanjutkan dengan penghitungan dari nilai *accuracy*, *precision*, *recall*, dan *F1-scores*. *Precision* (presisi) mengukur seberapa akurat model dalam mengklasifikasikan data sebagai positif. *Recall* (sensitivitas) mengukur seberapa baik model dalam menemukan kembali (mengklasifikasikan dengan benar) data yang sebenarnya positif. *F1-score* adalah ukuran yang menggabungkan *precision* dan *recall* untuk memberikan gambaran keseluruhan tentang kinerja model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{F1 - Scores} = \frac{2(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (4)$$

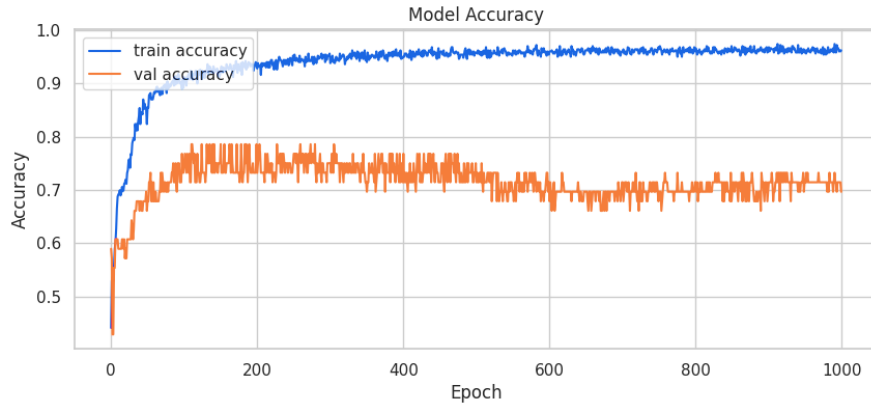
Di mana TP adalah *True Positive*, TN adalah *True Negative*, FP adalah *False Positive*, FN adalah *False Negative*.

2.6. Analisis

Proses analisis dilakukan terkait performa dari model yang dihasilkan berdasarkan parameter evaluasi yang digunakan. Selain itu pemanfaatan *word cloud* dari hasil praproses dataset juga ditampilkan untuk memberikan gambaran relevansi bidang permasalahan berdasarkan setiap kelas label, khususnya untuk label kelas positif dan label kelas negatif.

3. HASIL DAN PEMBAHASAN

Berdasarkan eksperimen yang telah dilaksanakan dapat diketahui bahwa secara umum *accuracy* model untuk tes dataset yang digunakan sebesar 76,26 %. Hal ini berarti model secara umum memiliki performa yang baik dalam mengklasifikasikan sentimen sesuai kelasnya namun masih ada beberapa sentimen yang diklasifikasikan tidak sesuai. Gambar 3 menampilkan hasil pelatihan model LSTM selama 1000 *epoch*.



Gambar 3. Proses Pelatihan Model LSTM

Setelah model klasifikasi berbasis LSTM dilatih dan diketahui *accuracy* secara keseluruhan, selanjutnya adalah mengetahui informasi terkait *accuracy*, *precision*, *recall*, *F1-scores* dari masing – masing kelas. *Confusion matrix* dapat digunakan untuk menghitung parameter evaluasi model ini. Gambar 4 menampilkan *confusion matrix* dari model yang sudah dibuat untuk klasifikasi sentimen kepuasan pelayanan puskesmas berbasis LSTM dan Tabel 4 menampilkan hasil penghitungan *parameter precision*, *recall*, dan *F1-scores*

Pada kelas negatif *precision* mengukur sejauh mana prediksi negatif yang dilakukan oleh model benar-benar merupakan sentimen negatif. Dalam kasus ini, *precision* untuk sentimen negatif adalah 0.790, yang menunjukkan bahwa sebagian besar prediksi negatif model untuk sentimen negatif adalah benar. *Recall* mengukur sejauh mana model mampu mendeteksi sentimen negatif secara keseluruhan. Dalam kasus ini, *recall* adalah 0.778, yang berarti model dapat menemukan kembali sebagian besar sentimen negatif yang ada dalam data. *F1-score* adalah ukuran kombinasi antara *precision* dan *recall*. *F1-score* untuk sentimen negatif adalah 0.784, mencerminkan keseimbangan yang baik antara *precision* dan *recall* untuk klasifikasi sentimen negatif.

Tabel 4. Hasil Penghitungan *Precision*, *Recall*, dan *F1-Scores* pada Setiap Kelas Sentimen

No	Kelas Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1-Scores</i>
1	Negatif	0.790	0.778	0.784
2	Netral	0.583	0.656	0.618
3	Positif	0.878	0.818	0.847

		Prediction		
		negative (0)	neutral (1)	positive (2)
Actual	negative (0)	49	11	3
	neutral (1)	9	21	2
	positive (2)	4	4	36

Gambar 4. *Confusion Matrix* Model Klasifikasi Sentimen Pelayanan Puskesmas

Pada kelas netral *precision* mengukur sejauh mana prediksi netral yang dilakukan oleh model benar-benar merupakan sentimen netral. Dalam kasus ini, *precision* untuk sentimen netral adalah 0.583, yang menunjukkan bahwa sebagian besar prediksi netral model untuk sentimen netral adalah benar. *Recall* mengukur sejauh mana model mampu mendeteksi sentimen netral secara keseluruhan. Dalam kasus ini, *recall* adalah 0.656, yang berarti model dapat menemukan kembali sebagian besar sentimen netral yang ada dalam data. *F1-score* adalah ukuran kombinasi antara *precision* dan *recall*. *F1-score* untuk sentimen netral adalah 0.618, mencerminkan keseimbangan antara *precision* dan *recall* untuk klasifikasi sentimen netral.

Pada kelas positif *precision* mengukur sejauh mana prediksi positif yang dilakukan oleh model benar-benar merupakan sentimen positif. Dalam kasus ini, *precision* untuk sentimen positif adalah 0.878, yang menunjukkan bahwa sebagian besar prediksi positif model untuk sentimen positif adalah benar. *Recall* mengukur sejauh mana model mampu mendeteksi sentimen positif secara keseluruhan. Dalam kasus ini, *recall* adalah 0.818, yang berarti model dapat menemukan kembali sebagian besar sentimen positif yang ada dalam data. *F1-score* adalah ukuran kombinasi antara *precision* dan *recall*. *F1-score* untuk sentimen positif adalah 0.847, mencerminkan keseimbangan antara *precision* dan *recall* untuk klasifikasi sentimen positif.

Berdasarkan Tabel 4 juga diketahui bahwa kelas netral adalah kelas yang paling sulit untuk diklasifikasikan, terlihat dari nilai *precision*, *recall*, dan *F1-scores* yang ada. Hal ini disebabkan oleh kelas yang tidak seimbang dimana kelas netral paling sedikit, kelas negatif memiliki rasio yang paling besar sehingga bisa lebih dominan terhadap netral. Selain itu adalah kompleksitas dan variabilitas dari teks kelas netral dimana tidak memiliki ciri khas tertentu dan sangat bergantung pada sudut pandang individu sehingga identifikasi kelas ini menjadi lebih sulit. Pada Tabel 5 menampilkan sampel dari hasil klasifikasi dengan model LSTM yang digunakan menunjukkan sampel kelas netral dari label asli, hasilnya kelas netral cenderung diklasifikasikan ke kelas positif dan dominan ke kelas negatif. Gambar 5 menampilkan hasil dari visualisasi kata yang sering muncul pada dataset dalam bentuk *word cloud* secara garis besar dapat terlihat bahwa beberapa kata yang sering muncul masuk ke dalam setiap kelas sentimen, ini mengarah pada variabilitas yang minim dari kata unik bagi setiap kelas.

Tabel 5. Sampel Hasil Klasifikasi Model pada Tes Dataset

No	Opini	Hasil Klasifikasi Model LSTM	Klasifikasi Label Asli
Klasifikasi Benar			
1	pelayanan sudah bagus tingkatkan pelayanan	positif	positif
2	loket pendaftaran kurang sopan	negatif	negatif
3	waktu tunggu cukup lama tapi wajar	netral	netral
Klasifikasi Salah			
1	kedepan agar pelayanan selalu ditingkatkan	positif	netral
2	ruang tunggu cukup padat dan antri	negatif	netral
3	menunggu pendaftaran lama tapi di poli cepat	negatif	netral



Gambar 5. *Word Cloud* dari Sentimen Dataset

4. KESIMPULAN

Penelitian ini bertujuan untuk mengeksplorasi potensi penerapan *deep learning* dalam kasus ini adalah arsitektur LSTM untuk klasifikasi sentimen kepuasan pelayanan Puskesmas. Hasil dari penelitian ini menunjukkan bahwa model LSTM dapat digunakan secara efektif untuk mengklasifikasikan sentimen dari masyarakat terhadap pelayanan Puskesmas. Akurasi keseluruhan dari model ini adalah 76.26%, yang menunjukkan bahwa model berhasil melakukan prediksi dengan cukup baik. Pelayanan puskesmas yang standar diseluruh Indonesia sudah diatur oleh kementerian kesehatan Republik Indonesia, yang berarti hal yang dievaluasi melalui survei kepuasan pelayanan Puskesmas ini akan memiliki domain yang sama. Oleh karena itu hasil dari penelitian ini dengan memanfaatkan laporan survei sebagai dataset untuk menganalisa sentimen pelayanan dapat digunakan sebagai bahan pembandingan untuk potensi penerapan dan integrasi *deep learning* dalam pengembangan sistem kesehatan digital Indonesia khususnya yang berkaitan dengan pelayanan Puskesmas. Namun, perlu diperhatikan bahwa performa model memiliki perbedaan untuk masing-masing kelas sentimen (negatif, netral, dan positif). Sentimen netral menunjukkan performa yang sedikit lebih rendah dibandingkan dengan sentimen positif dan negatif. Hal ini dapat disebabkan oleh beberapa faktor, seperti ketidakseimbangan dalam jumlah sampel untuk masing-masing kelas, kompleksitas sentimen netral, variabilitas dalam teks, dan perbedaan subyektivitas dalam penilaian sentimen netral. Kedepan dapat dilakukan penambahan dataset dan memastikan jumlah kelas bisa seimbang, selain itu dapat dilakukan proses analisa dengan menggunakan teknik *machine learning* maupun *deep learning* lainnya.

DAFTAR PUSTAKA

- [1] F. E. B. Setyawan, S. Supriyanto, E. Ernawaty, and R. Lestari, "Understanding patient satisfaction and loyalty in public and private primary health care," *J Public Health Res*, vol. 9, no. 2, p. jphr-2020, 2020.
- [2] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," *J Big Data*, vol. 6, no. 1, pp. 1–25, 2019.
- [3] L. Rajabion, A. A. Shaltooki, M. Taghikhah, A. Ghasemi, and A. Badfar, "Healthcare big data processing mechanisms: The role of cloud computing," *Int J Inf Manage*, vol. 49, pp. 271–289, 2019.
- [4] Z. Lv and L. Qiao, "Analysis of healthcare big data," *Future Generation Computer Systems*, vol. 109, pp. 103–110, 2020.
- [5] M. Z. Yumarlin, J. E. Bororing, and S. Rahayu, "Analisis Sentimen Terhadap Layanan Tokopedia Berdasarkan Twitter dengan Metode Klasifikasi Support Vector Machine," *Smart Comp: Jurnalnya Orang Pintar Komputer*, vol. 12, no. 1, pp. 153–163, 2023.
- [6] S. K. Banbhrani, B. Xu, H. Lin, and D. K. Sajjani, "Spider Taylor-ChOA: Optimized Deep Learning Based Sentiment Classification for Review Rating Prediction," *Applied Sciences*, vol. 12, no. 7, p. 3211, 2022.
- [7] G. Indrawan, H. Setiawan, and A. Gunadi, "Multi-class SVM Classification Comparison for Health Service Satisfaction Survey Data in Bahasa," *HighTech and Innovation Journal*, vol. 3, no. 4, pp. 425–442, 2022.
- [8] M. H. Setiawan, I. G. A. Gunadi, and G. Indrawan, "Klasifikasi Pelayanan Kesehatan Berdasarkan Data Sentimen Pelayanan Kesehatan menggunakan Multiclass Support Vector Machine," *Jurnal Sistem dan Informatika (JSI)*, vol. 17, no. 1, pp. 47–54, 2022.
- [9] G. S. N. Murthy, S. R. Allu, B. Andhavarapu, M. Bagadi, and M. Belusonti, "Text based sentiment analysis using LSTM," *Int. J. Eng. Res. Tech. Res*, vol. 9, no. 05, 2020.
- [10] X. Huang *et al.*, "Lstm based sentiment analysis for cryptocurrency prediction," in *Database Systems for Advanced Applications: 26th International Conference, DASFAA 2021, Taipei, Taiwan, April 11–14, 2021, Proceedings, Part III* 26, Springer, 2021, pp. 617–621.
- [11] C. for P. H. I. F. K. U. Udayana, "Laporan Survei Kepuasan Masyarakat terhadap Puskesmas di Denpasar dan Gianyar," 2021.
- [12] U. Hasanah, T. Astuti, R. Wahyudi, Z. Rifai, and R. A. Pambudi, "An experimental study of text preprocessing techniques for automatic short answer grading in Indonesian," in *2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE)*, IEEE, 2018, pp. 230–234.
- [13] M. A. Rosid, A. S. Fitrani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, "Improving text preprocessing for student complaint document classification using sastrawi," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 2020, p. 012017.
- [14] B. Siswanto and Y. Dani, "Sentiment Analysis about Oximeter as Covid-19 Detection Tools on Twitter Using Sastrawi Library," in *2021 8th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, IEEE, 2021, pp. 161–164.
- [15] M. A. Rosid, A. S. Fitrani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, "Improving text preprocessing for student complaint document classification using sastrawi," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 2020, p. 012017.