

Prediksi Penyakit Parkinson Melalui Dataset Rekam Suara Dengan Menggunakan Algoritma Deep Neural Network

¹Kristina Ronalita Nainggolan, ^{2*}Fanindia Purnamasari, ³Annisa Fadhillah Pulungan
^{1, 2, 3}Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara

¹kristinarn1402@gmail.com, ^{2*}fanindia@usu.ac.id, ³annisafpulungan@usu.ac.id

ABSTRAK

Penyakit Parkinson merupakan suatu gangguan sistem saraf kronis dan progresif yang dapat memengaruhi gerakan tubuh manusia. Pemeriksaan terhadap penyakit parkinson belum dilakukan seperti memeriksa keadaan medis pada umumnya, sehingga banyak pasien tidak menyadari bahwa mereka ternyata memiliki gejala penyakit Parkinson. Melalui pendekatan teknologi, penyakit parkinson dapat diperiksa melalui data rekam suara dari seorang pasien. Data rekam suara yang diperoleh dari dataset berupa data numerik yang kemudian perlu diproses. Data tersebut diproses dengan menggunakan pendekatan komputasi. Penelitian ini mengusulkan algoritma Deep Neural Network. Pada dataset yang diperoleh ditemukan ketidakseimbangan data sehingga diusulkan metode SMOTE (*Synthetic Minority Oversampling Technique*) yang digunakan sebagai teknik oversampling data. Selanjutnya diterapkan algoritma Deep Neural Network untuk memprediksi parkinson pada data tersebut dengan menggunakan parameter pada model seperti penggunaan aktivasi ReLU pada *hidden layer*, aktivasi sigmoid pada *output layer*, *dropout*, epochs 200, batch size 16, optimizer Adam dengan penggunaan learning rate sebesar 0.001, dan metrics accuracy, sehingga memberikan hasil akurasi sebesar 97%.

Kata Kunci: prediksi, penyakit parkinson, deep neural network, oversampling

PENDAHULUAN

Teknologi telah berkembang pada berbagai bidang, salah satunya adalah pada bidang kesehatan. Kecakapan suatu teknologi pada bidang kesehatan memberikan dampak positif dalam pemeriksaan medis serta mampu memberikan hasil diagnosis yang akurat, proses penyembuhan atas suatu penyakit, dan memudahkan proses sistem administrasi lainnya. Hasil diagnosis pada pasien sangat membantu pasien dalam melakukan proses penyembuhan pada tahap selanjutnya, namun hal ini berbeda dengan pasien penyakit Parkinson. Penyakit Parkinson adalah gangguan *neurodegenerative*, yaitu kelainan yang mempengaruhi sebagian besar sistem syaraf progresif yang menyebabkan gangguan pada sistem motorik manusia. Gejala pada penyakit ini umumnya meliputi tremor, gerakan yang melambat, otot yang terasa kaku, gangguan postur tubuh dan keseimbangan, kehilangan pergerakan otomatis, perubahan berbicara dan perubahan dalam gerakan menulis. Walaupun penyebab penyakit ini tidak diketahui secara pasti, beberapa peneliti menemukan bahwa adanya perubahan mutasi genetik yang spesifik dan dapat juga disebabkan paparan oleh racun tertentu yang tanpa disadari berdampak kepada seseorang (Clinic M, 2019). Ramaker et al., (2002) menyebutkan beberapa skala telah dikembangkan untuk memeriksa tahapan penyakit parkinson tersebut. Sampai saat ini belum ada metode pemeriksaan yang dapat memastikan diagnosis penyakit parkinson sehingga beberapa pemeriksaan alternatif dilakukan seperti melalui tes darah, single photon emission computed tomography (SPECT), pemindaian dengan MRI, CT scan, dan PET scan (Clinic M., 2019), EEG voltage atau citra MRI, citra tulisan tangan, data sensory dan

data rekam sinyal suara pasien (Shaban, 2023).

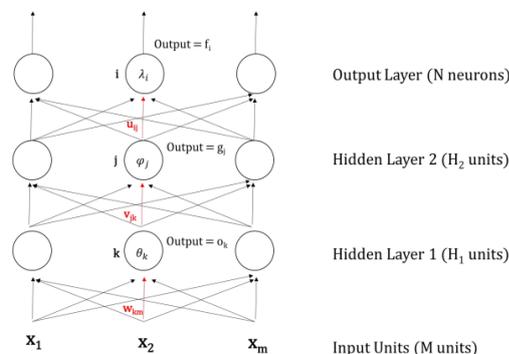
Diagnosis penyakit Parkinson selama ini didasarkan pada riwayat medis seseorang dan pemeriksaan neurologis secara berkala. Hal ini disebabkan karena penyakit Parkinson adalah penyakit yang menyerang seseorang secara perlahan, sebagaimana gejala awal yang tidak terlihat menyebabkan pasien berpikir bahwa mereka menderita penyakit biasa seperti kelelahan. Selain itu, tidak tersedianya tes medis seperti pemeriksaan fisik yang dapat mendiagnosis penyakit parkinson tersebut. Salah satu cara untuk mengetahui bahwa seseorang tersebut memiliki penyakit Parkinson atau tidak adalah dengan melakukan prediksi melalui data berupa rekam medis suara pasien yang diubah ke dalam bentuk ASCII berdasarkan variabel yang telah ditentukan (Little,2008).

Oleh karena itu, data berupa citra MRI, tulisan tangan, data sensori dan data rekam sinyal suara pasien memerlukan satu pendekatan teknologi dan komputasi agar data tersebut dapat dibaca dan dipahami. Hal ini seperti pada pendekatan machine learning telah dilakukan pada beberapa penelitian sebelumnya, antara lain Prakoso, et al. (2020) mengusulkan sebuah sistem klasifikasi pada penyakit Parkinson dengan menggunakan metode K-Nearest Neighbor yang memberikan akurasi sebesar 90.89%. Soumaya, et al. (2020) mendeteksi penyakit Parkinson dengan menggabungkan dua metode yaitu algoritma genetika dan Support Vector Machine Classifier dengan menggunakan fitur akustik and dekomposisional dari suara yang mampu memberikan hasil akurasi sebesar 91,18%. Abdurrahman dan Sintawati, (2020) mengusulkan XGBoost dengan menggunakan feature selection dan tanpa menggunakan feature selection 85,60% dan 84,80%. Asmae et al., (2020) mengusulkan KNN dan ANN untuk mengidentifikasi penyakit parkinson dari suara pasien yang memberikan. Hasil akurasi 79%. Ashour, El-Attar et al.(2020) menggunakan LSTM pada data sinyal FOG dari beberapa sensor yang terpasang di tubuh pasien selama proses pengambilan. Karabayir et.al (2020) mengusulkan gradient boosting dengan menggunakan dataset rekaman suara. Purba (2022) mengusulkan mendiagnosis seorang pasien Parkinson dengan memanfaatkan visi komputer untuk mendeteksi gerakan tangan pasien dan non pasien melalui tekanan pada gambar pola yang diberikan. Berdasarkan latar belakang dan beberapa penelitian yang telah dilakukan, penulis berharap dari tujuan penelitian ini untuk dapat melakukan prediksi awal dalam diagnosis penyakit Parkinson melalui faktor yang dikaitkan dalam menganalisis pola penyakit. Dari penelitian ini diharapkan penggunaan metode yang telah dipilih dapat memberikan hasil akurasi yang akurat dan dapat memberikan pernyataan bahwa seseorang menderita penyakit Parkinson atau tidak.

TINJAUAN PUSTAKA

Arsitektur Deep Neural Network (DNN)

Algoritma DNN ini merupakan pembelajaran lebih lanjut dari algoritma Artificial Neural Network (ANN) yang banyak digunakan dalam pengambilan keputusan. Jika biasanya ANN memiliki satu *input layer*, satu *output layer*, dan maksimal satu *hidden layer* berbeda dengan DNN yang harus memiliki lebih dari satu *hidden layer* (Ashraf, et al. 2019)..



Gambar 1. Arsitektur DNN (Wira Gotama Putra, 2017)

Oversampling Data

Sampling adalah teknik penyeimbangan data. Teknik sampling dibagi menjadi dua yaitu

undersampling dan *oversampling*. *Undersampling* adalah teknik menyeimbangkan dataset dengan mengurangi ukuran data yang berlimpah sedangkan *oversampling* adalah teknik menyeimbangkan data dengan meningkatkan ukuran pada kelas yang minoritas. Pada penelitian ini teknik sampling yang digunakan adalah *oversampling*. Teknik *oversampling* berfungsi dalam menambahkan data pada kelas minoritas untuk menyeimbangkan distribusi data pada kelas. Pendekatan *oversampling* lebih banyak digunakan dibandingkan *undersampling* karena metode *undersampling* dilakukan untuk mengeliminasi data pada kelas mayoritas sehingga menyebabkan hilangnya informasi penting dari data tersebut. Penelitian yang dilakukan oleh (Chawla, et al. 2022) mengusulkan teknik baru dari *oversampling* yaitu *SMOTE (Synthetic Minority Over-sampling Technique)*, yaitu menghasilkan data sintetis berdasarkan jarak antara data minoritas dengan data minoritas terdekat sehingga akan terbentuk data sintetis baru antara dua data minoritas. Berdasarkan penelitian yang dilakukan oleh (Chawla, et al. 2002), rumus yang digunakan dalam menyatakan data sintetis pada *SMOTE* adalah:

$$D_{new} = D_i + (\widehat{D}_i - D_i) \times \delta \quad (1)$$

Dimana:

D_{new} = data sintetis

D_i = contoh data minoritas

\widehat{D}_i = salah satu k-nearest neighbors dari D_i (nilai terdekat)

δ = angka acak antara 0 dan 1

Penyakit Parkinson

Penyakit Parkinson atau sering disebut sebagai tremor merupakan suatu bentuk kelainan pada syaraf yang dapat ditandai dengan hilangnya sel dan gliolisis pada pasangan inti batang otak manusia yang disebut *substantia nigra* yaitu pigmen nukleus yang mengandung sel kaya zat dopamine yang berperan dalam mengatur pergerakan tubuh (Balestrino & Schapira, 2019). Gejala umum yang sering dialami oleh pasien adalah tremor atau gerakan gemetar yang terjadi secara tidak terkendali ditambah gejala-gejala lain yang menjadi ciri umum pada penyakit Parkinson. Terdapat dua jenis gejala yang umumnya terjadi pada pasien penderita penyakit Parkinson yaitu gejala motorik antara lain gerakan tremor pasien, gerakan *bradykinesia* (melambat), kekakuan pada otot-otot gerak, gangguan postur dan keseimbangan, akinesia dan kecepatan dalam berbicara. Sedangkan gejala non-motorik meliputi disfungsi otonom, depresi, disfungsi kognitif, psikosis, dan mengalami gangguan tidur (Chaudhuri & Schapira, 2009). Penyakit parkinson dapat diamati dari getaran suara pasien dengan merekam fonasi (proses bersuara dimana pita suara di tenggorokan menghasilkan bunyian atau tanpa suara) seperti yang dilakukan pada penelitian Little et al., (2008). Hasil rekam suara pasien tersebut menemukan beberapa variabel dalam penentu atau memprediksi seseorang terjangkit parkinson atau bukan. Berdasarkan tes suara pasien tersebut, para ahli akhirnya menentukan beberapa variabel yang menjadi bahan pertimbangan diagnosis penyakit Parkinson. Menurut (Tsanas, 2009),(Little, et al. 2008) beberapa hasil rekaman tes suara pasien sebagai berikut:

Tabel 1. Daftar ukuran yang diaplikasikan pada sinyal suara rekam pasien

No	Variabel	Keterangan
1	MDVP:Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz)	dan MDVP:Fo(Hz) adalah rata-rata frekuensi dasar suara, MDVP:Fhi(Hz) merupakan nilai maksimum dari frekuensi dasar suara, dan MDVP:Flo(Hz) merupakan nilai minimum dari frekuensi dasar. Frekuensi dasar ini dihitung dalam Hertz (Hz) dimana pada pria rata-rata yang dihasilkan sebesar 80-200 Hz dan pada wanita sebesar 200-300 Hz
2	Jitter	Tingkat variasi perubahan suara dari satu siklus ke siklus lain dalam frekuensi dasar
3	Shimmer	Tingkat variasi amplitudo

No	Variabel	Keterangan
4	HNR (<i>Harmonics to Noise ratio</i>) dan NHR (<i>Noise to Harmonics ratio</i>)	Ukuran rasio kebisingan terhadap komponen tonal dalam suara
5	RPDE (<i>Recurrence Period Density Entropy</i>)	Kemampuan pita suara yang stabil untuk mempertahankan getaran dan mengukur distorsi dari frekuensi dasar.
6	D2	Dimensi korelasi
7	DFA (<i>Detrended Fluctuation Analysis</i>)	Tingkat kebisingan turbulen dalam sinyal suara dan mengukur kesamaan diri stokastik dari kebisingan yang disebabkan aliran udara
8	PPE (<i>Pitch period entropy</i>)	Suatu bentuk perubahan dalam ucapan suara yang dipengaruhi lingkungan akustik

Rajeswari et al., (2022) menggunakan fitur dari audio sinyal seperti Harmonic/NoiseRatio, Jitter, Noise/Harmonic Ratio, Shimmer dengan menggunakan algoritma pra proses sinyal suara seperti Mel Frequency Cepstrum Coefficient(MFCC), Time Fre-quency Features, dan Vocal Fold Feature.

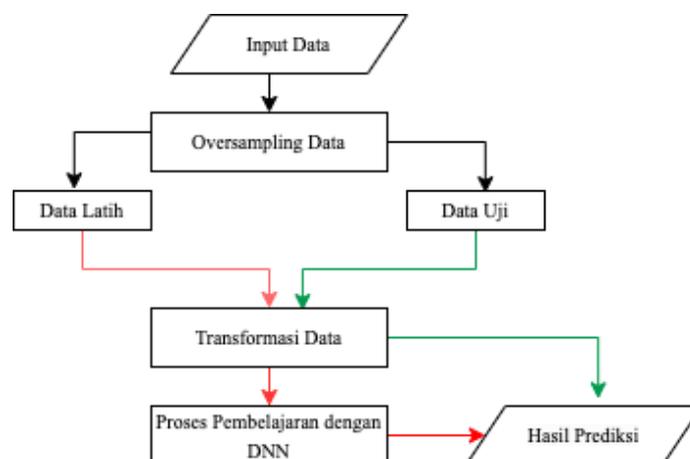
METODE PENELITIAN

Data Penelitian

Penelitian ini menggunakan dataset yang diperoleh dari situs web UCI Machine Learning. Dataset ini berasal dari penelitian yang telah dilakukan berupa data rekam suara pasien yang dilakukan setiap enam bulan sekali berjumlah 31 pasien dengan total 195 rekaman suara (Little, 2008). Data ini terdiri dari 22 parameter penentu diagnosis penyakit Parkinson (Little, et al. 2007). Data yang digunakan untuk mengimplementasikan algoritma deep neural network adalah sejumlah 146 data latih dan 49 data uji.

Rancangan Penelitian

Penelitian ini dimulai dengan memeriksa keseluruhan dataset yang telah diperoleh. Melihat adanya ketidakseimbangan data antara data positif dan negatif pasien yang terdiagnosis parkinson, peneliti mengusulkan teknik oversampling data dengan teknik SMOTE kemudian membagi menjadi data latih dan data uji. Data yang diperoleh melewati praproses dengan melakukan transformasi data. Tahapan selanjutnya tahap pembelajaran bagi data latih dengan menggunakan deep neural network untuk menghasilkan output dengan dua kelas yakni positif atau negatif parkinson. Gambar 2 menunjukkan rancangan penelitian.



Gambar 2 Rancangan Penelitian

1. Input Data

Data yang diinput berupa data numerik dari data rekam suara .csv yang terdiri dari 22 atribut.

2. Oversampling Data

Pada penelitian ini, dalam menyeimbangkan data digunakan metode SMOTE yang digabungkan dengan menggunakan teknik pra pemrosesan *Tomek Links* sebagai teknik pembersihan data dari data sintetik yang telah dibentuk sebelumnya oleh SMOTE. Hasil kinerja model menjadi lebih baik dengan menggunakan gabungan SMOTE beserta Tomek Links daripada hanya menggunakan SMOTE dalam mengatasi data klasifikasi yang tidak seimbang (Zeng, et al. 2016).

3. Transformasi Data

Transformasi data yang digunakan adalah *standard scaler*, yakni teknik melakukan skala ulang pada parameter untuk memastikan rata-rata nilai menjadi 0 dan menggunakan standard deviasi (nilai akar kuadrat dari varians dan menunjukkan standard penyimpangan data terhadap nilai rata-ratanya) menjadi 1 sehingga parameter berbentuk menjadi distribusi normal sehingga bobot mudah untuk dipelajari. Pada tahapan ini terdapat dua fungsi yang digunakan dalam menyatakan data dalam pelatihan dan pengujian data, yaitu fungsi *fit_transform()* merupakan fungsi yang digunakan pada data latih dan juga berfungsi untuk mempelajari parameter penskalaan data tersebut. Model akan mempelajari rata-rata dan standard deviasi dari fitur training set. Parameter yang telah dilatih kemudian akan digunakan kembali dalam penskalaan data uji. Sementara fungsi *transform()* akan diterapkan pada data uji dengan menggunakan hasil rata-rata dan standard deviasi yang telah dipelajari pada *fit_transform()* sebelumnya. Secara singkatnya, nilai rata-rata dan standard deviasi yang telah dilakukan pada data latih akan disimpan dan digunakan kembali pada data uji.

```
array([[1.482720e+02, 1.649890e+02, 1.422990e+02, ..., 8.784000e-02,
        2.344336e+00, 1.864890e-01],
       [1.166760e+02, 1.378710e+02, 1.113660e+02, ..., 3.341470e-01,
        2.405554e+00, 3.689750e-01],
       [1.564050e+02, 1.893980e+02, 1.428220e+02, ..., 3.401760e-01,
        2.856676e+00, 3.221110e-01],
       ...,
       [2.288320e+02, 2.346190e+02, 2.236340e+02, ..., 1.868900e-02,
        2.498224e+00, 9.353400e-02],
       [1.697740e+02, 1.917590e+02, 1.514510e+02, ..., 4.147580e-01,
        3.413649e+00, 4.575330e-01],
       [2.289690e+02, 2.395410e+02, 1.132010e+02, ..., 6.274000e-03,
        2.118596e+00, 9.154600e-02]])
```

Gsmbsr 3. Array Data Uji sebelum dilakukan standarisasi

```
[[ 1.35911368  0.60259556 -0.14910966 ... -1.88523313  0.20587358
 -1.17083771]
 [-0.18299311  0.10559502 -0.8223841 ... -1.16903589 -1.64739932
  0.14554857]
 [ 0.06250971  0.47500454 -0.83601731 ...  0.61498329 -0.37837758
 -1.09483052]
 ...
 [-0.86984174 -0.73985484 -0.09813739 ...  0.98372688 -0.11027695
  1.33939797]
 [ 0.56421874 -0.05220597  1.12016092 ... -0.36512928  0.40787792
 -0.39206272]
 [-1.30082165 -1.10335865 -0.65449885 ... -1.75354571 -0.88019042
 -0.34799853]]
```

Gsmbsr 4. Array Data Uji setelah dilakukan standarisasi

4. Proses Pembelajaran dengan DNN

Data yang telah melewati tahap praproses, kemudian lanjut ke tahap proses pembelajaran atau melatih model dengan menggunakan dataset yang tersedia. Pada tahapan ini dibentuk sebuah model menggunakan metode Deep Neural Network dimana model akan dibentuk dengan melakukan beberapa pengaturan hyperparameter untuk memperoleh parameter yang mendapat akurasi yang tepat.

5. Hasil Prediksi

Pada tahapan ini, model akan memprediksi hasil output dengan dua kelas yaitu positif dan negatif parkinson

HASIL DAN PEMBAHASAN

Hasil Oversampling data

Hasil dari penyeimbangan data menggunakan penggabungan SMOTE dan Tomek Link didapatkan perbandingan data positif dan negatif 146:146, dimana data negatif yang menjadi minoritas diberikan data sintetik dalam menyeimbangkan data dengan data positif yang diberikan. Total dari seluruh data yang digunakan sebanyak 292 data yang telah dibagi menjadi 75% data training yaitu sebanyak 219 data dan 25% untuk data testing sebanyak 73 data.

Tabel 2. Perbandingan data setelah oversampling

Data setelah oversampling	
Data latih	219
Data uji	73
Total	292

Tabel 3. Perbandingan Data Latih dan data uji setelah oversampling

Data Latih		Data Uji	
Positif	Negatif	Positif	Negatif
105	114	42	31
Total	219	Total	73

Hasil Pembelajaran dengan Arsitektur DNN

Model arsitektur DNN yang digunakan untuk melakukan prediksi terhadap data uji yang dimiliki. Node pada *input layer* merupakan 22 atribut dari dataset pasien penyakit Parkinson. Pada lapisan *hidden layer* berisi neuron dan fungsi aktivasi sedangkan *output layer* terdapat dua node yaitu hasil prediksi (positif atau negatif). Model terbaik untuk memprediksi penyakit Parkinson dilatih dengan menggunakan parameter seperti pada Tabel 4.

Tabel 4. Parameter yang digunakan

Parameter	Nilai
Aktivasi Input Layer	ReLU
Aktivasi Hidden Layer	ReLU
Aktivasi Output Layer	Sigmoid
Dropout	0,2 (input) dan 0,5 (hidden)
Epochs	200
Batch Size	16
Loss	Binary cross-entropy
Optimizer	Adam
Metrics	Accuracy

Aktivasi pada *input layer* terdiri dari 64 unit (mewakili jumlah neuron pada lapisan) dengan menggunakan aktivasi ReLU dan pada *hidden layer* dengan masing-masing nilai 128, 256, 512 dan pada output layer terdiri dari satu unit. Fungsi ReLU ini hanya dapat digunakan pada hidden layer. Sedangkan pada lapisan output, aktivasi yang digunakan adalah sigmoid karena fungsi sigmoid dapat memberikan hasil yang lebih baik untuk kasus klasifikasi (Sharma, et al. 2020). Setiap lapisan berisi dense layer dan dropout layer. Dense layer digunakan untuk menerima setiap input dari neuron pada lapisan sebelumnya. Sedangkan penggunaan dropout pada model ini adalah untuk

mencegah terjadinya overfitting pada model. Pada model, nilai dropout yang digunakan pada input layer dan hidden layer adalah 0.2. Sedangkan pada hidden layer, nilai dropout yang digunakan adalah 0.5.

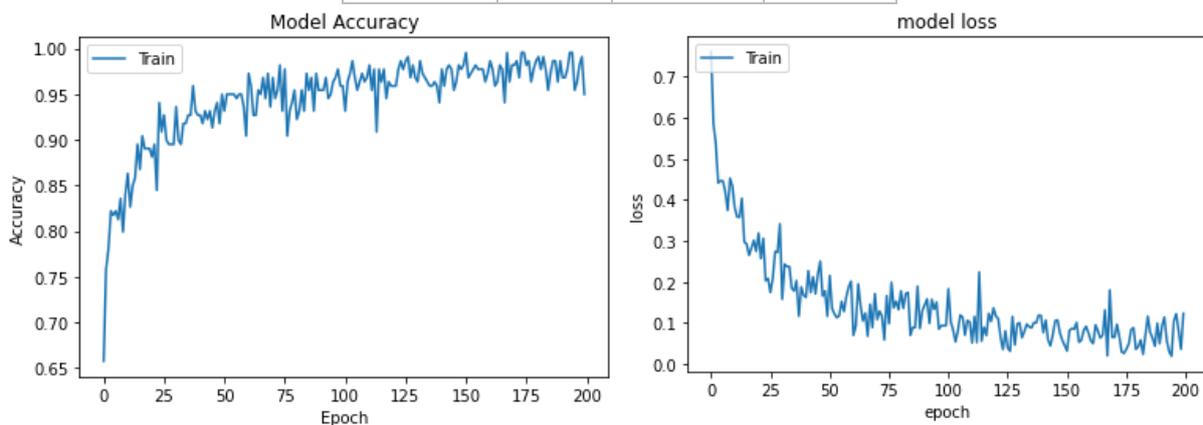
Berdasarkan penelitian yang dilakukan oleh Srivastava, et al. (2014) penggunaan nilai dropout sebesar 0.5 merupakan nilai optimal dalam meningkatkan hasil terbaik pada proses pelatihan model. Selanjutnya, model yang diuji menggunakan loss berupa binary cross entropy yaitu mengklasifikasikan nilai 1 (positif) dan 0 (negatif) dan disesuaikan dengan penggunaan fungsi aktivasi Sigmoid pada output layer yang bernilai 1 dan 0. Optimizer yang digunakan adalah Adam dan metrics berupa *accuracy*. Selain menggunakan optimizer Adam, pada penelitian ini juga menggunakan learning rate yang berfungsi menghitung nilai koreksi bobot pada proses pelatihan sehingga dapat memberikan hasil loss minimum.

Pembahasan

Penelitian ini mencoba beberapa learning rate dalam proses pelatihan yaitu 0.01, 0.001, 0.0001. Dari beberapa hasil pelatihan learning rate 0.001 dapat memberikan hasil akurasi yang baik dan nilai loss yang rendah. Epochs yang digunakan sebesar 200 dan batch size sebesar 16 untuk skala dataset yang digunakan. Nilai epochs 200 berdasarkan hasil pelatihan data dengan membandingkan epoch 100, 200, 300, dan 400 pada pembagian data 75:25%. Pada pengujian, epoch 200 dan batch size 400 pada pembagian data 75:25 memberikan hasil akurasi yang cukup baik seperti pada Tabel 5. Penggunaan epoch tersebut sebagai perulangan atas seluruh dataset yang lewat menuju jaringan syaraf. Pada pelatihan model, akurasi terbaik didapatkan pada epoch ke-152 dimana model akurasi yang berjalan pada proses training sebesar 99,54.

Tabel 5. Akurasi Tiap Epoch

Batch size	LR	Epoch	Akurasi
16	0.001	100	91,78
		200	97,26
		300	95,89
		400	95,89



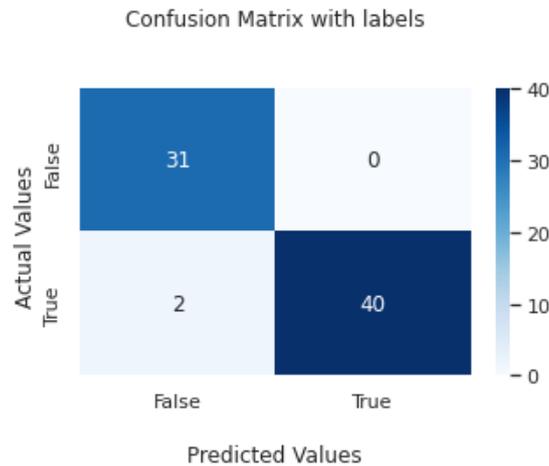
Gambar 3 Hasil Akurasi dan Loss dari model yang dilatih

Gambar 3 menampilkan plot atau pergerakan epoch loss selama pelatihan. Pergerakan loss diawali dari nilai 1 dan secara drastis turun pada kisaran 0,2 sampai 0,4 pada epoch 1-25. Namun, setelah itu epoch tetap berjalan walau pergerakannya perlahan turun. Hal ini dikarenakan terdapat nilai loss atau kesalahan prediksi pada model dengan memperhitungkan probabilitas atau

ketidakpastian prediksi berdasarkan seberapa besar variasi prediksi dari nilai aktualnya sehingga nilai loss diharapkan bernilai minimum.

Hasil Akurasi

Berdasarkan hasil testing yang telah dilakukan terhadap data disimpulkan hasil dalam *heatmap confusion matrix*.



Gambar 4. Hasil Confusion Matrix

Hasil *confusion matrix* menyebutkan bahwa terdapat nilai True Positive yaitu pasien yang memiliki penyakit Parkinson sebanyak 40 pasien, False Positive menyatakan bahwa pasien menderita penyakit Parkinson namun faktanya tidak ada, True Negative menyatakan bahwa pasien tidak menderita penyakit Parkinson sedangkan faktanya benar sebanyak 31 pasien diprediksi mengidap penyakit parkinson, dan False Negative yang menyatakan bahwa pasien tidak menderita penyakit Parkinson namun faktanya seseorang tersebut menderita penyakit Parkinson sebanyak dua pasien. Model ini menghasilkan persentase Nilai Recall sebesar 95%, Precision 100%, F1-score 98%, dan nilai akurasi 97%.

KESIMPULAN

Deep neural network (DNN) digunakan untuk memprediksi data rekam suara pasien positif dan negatif penyakit parkinson. Model telah melakukan beberapa penggunaan hyperparameter tuning untuk memperoleh akurasi terbaik dengan menggunakan epoch sebesar 200 dan batch size 16. Prediksi akurasi tertinggi sebesar 97% dengan perbandingan jumlah data latih dan data uji sebesar 75:25. Pada penelitian ini menggunakan oversampling yaitu SMOTE dan Tomek Links dan penggunaannya pada data yang tidak seimbang berpengaruh baik serta mampu meningkatkan hasil akurasi prediksi. Pada penelitian selanjutnya diharapkan dilakukan pelatihan model dengan membandingkan performa model DNN menggunakan beragam batch size pada setiap epoch dan menggunakan data dari rumah sakit untuk membandingkan hasil yang lebih baik dari tiap dataset yang digunakan. Selain itu penggunaan metode lain dalam melakukan prediksi terkhusus penggunaan metode deep learning juga disarankan untuk membandingkan performa antar metode dalam melakukan prediksi pada penyakit Parkinson

REFERENSI

- Abdurrahman, G, and M Sintawati. 2020. Implementation of xgboost for classification of parkinson's disease. *Journal of Physics: Conference Series* 1-7.
- Amira S. Ashour, Amira El-Attar, Nilanjan Dey, Hatem Abd El-Kader, Mostafa M. Abd El-Naby, Long short term memory based patient-dependent model for FOG detection in Parkinson's disease, *Pattern Recognition Letters*, <https://doi.org/10.1016/j.patrec.2019.11.036>.

- Ashraf, Mohd, M., A. Rizvi, and Himanshu Sharma. 2019. Improved Heart Disease Prediction Using Deep Neural Network. *Asian Journal of Computer Science and Technology* 49-54.
- Asmae, Ouhmida, Raihani Abdelhadi, Cherradi Bouchaib, Sandabad Sara, and Khalili Tajeddine. 2020. Parkinson's Disease Identification using KNN and ANN Algorithms based on Voice Disorder. *IEEE Explore* 1-6.
- Balestrino, R, and A.H.V Schapira. 2019. Parkinson Disease. *European Journal of Neurology* 27-42.
- Chaudhuri, K Ray, and Anthony, H.V Schapira. 2009. Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *Lancet Neurol* 464-474.
- Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 16 321-357.
- Clinic M. Parkinson's disease (2019). <https://www.mayoclinic.org>.
- Karabayir I, Goldman SM, Pappu S, Akbilgic O. Gradient boosting for Parkinson's disease diagnosis from voice recordings. *BMC Med Inform Decis Mak*. 2020 Sep 15;20(1):228. doi: 10.1186/s12911-020-01250-7. PMID: 32933493; PMCID: PMC7493334.
- Little, Max A. 2008. UCI Machine Learning Repository. June 26. Accessed February 2022. <https://archive.ics.uci.edu/>.
- Little, Max A, Patrick E McSharry, Stephen J Roberts, Declan AE Costello, and Moroz Irene M. 2007. "Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection." *Nature Precedings* 1-35.
- Little, Max A., Member IEEE, Patrick E. McSharry, Senior Member IEEE, Eric J. Hunter, Jennifer Spielman, and Lorraine O. Ramig. 2008. Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease. *Nature Precedings* 1-27.
- Prakoso, Riyan Dwi Yulian, Bambang Soejono Wiriaatmadja, and Ferry Wahyu Wibowo. 2020. Sistem Klasifikasi Pada Penyakit Parkinson Dengan Menggunakan Metode K-Nearest Neighbor. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)* 63-68.
- Purba, Petrus S. dan Zein, A. 2022. Mendeteksi Penyakit Parkinson Dengan Opencv, Computer Vision, dan Spiral / Wave Test. – *Sainstech Vol. 32 No. 2*, hlm 76-81.
- Rajeswari , S., S., and Nair, M., Prediction of Parkinson's disease from Voice Signals Using Machine Learning. 2022. *Journal of Pharmaceutical Negative Results*, 2031-2035. <https://doi.org/10.47750/pnr.2022.13.S07.294>.
- Tsanas, Athanasios. 2009. Accurate Telemonitoring of Parkinson's Disease Progression by Non-invasive Speech Tests. *IEEE Transactions on Biomedical Engineering* 1-10.
- Shaban, M. 2023. Deep Learning for Parkinson's Disease Diagnosis: A Short Survey. *Computers*, 12(3), 58. <https://doi.org/10.3390/computers12030058>.
- Soumaya, Z, Belhoussine, D T, Nsiri Benayad, Korkmaz Yunus, and Ammoumou Abdelkrim. 2021. The Detection of Parkinson Disease using the Genetic Algorithm and SVM Classifier.
- Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* 15 1929-1958.
- Zeng, Ming, Beiji Zou, Faran Wei, Xiyao Liu, and Lei Wang. 2016. Effective prediction of three common diseases by combining SMOTE with Tomek links technique for imbalanced medical data." *IEEE Proceedings of ICOACS2016* 225-228.