

Implementasi Algoritma Artificial Neural Network dengan Aktivasi ReLU: Klasifikasi Tiroid

Lestari Yusuf¹, Taufik Hidayatulloh²

¹Universitas Nusa Mandiri
e-mail: lestari.lyf@nusamandiri.ac.id

²Universitas Bina Sarana Informatika
e-mail: taufik.tho@bsi.ac.id

Abstrak

Kelenjar tiroid yang memiliki peran penting dalam tubuh sebagai zat yang mempengaruhi metabolisme termasuk kedalam bidang endokrinologi. Terdapat 22.368 kasus kematian di Amerika Serikat yang diakibatkan kanker tiroid. Angka kematian tersebut harus terus ditekan agar tingkat kematian yang disebabkan oleh kelainan tiroid bisa berkurang. Satu diantara cara penekanan angka kematian tersebut bisa menggunakan perkembangan teknologi. Algoritma data mining yang menjadi metode pembelajaran sebuah data dalam mengenali pola yang kompleks dari bigdata dan juga beragam dalam pengambilan keputusan dalam dunia klinis adalah algoritma artificial neural network. Pada penelitian ini akan membuat klasifikasi penderita tiroid yang bisa kambuh atau tidak dan mengetahui faktor apa saja yang berhubungan dengan kesembuhan pasien kelainan tiroid penelitian ini menggunakan model algoritma Artificial neural network dengan aktivasi ReLU. Penelitian ini menggunakan thyroid yang merupakan data sekunder dari kaggle.com. sebanyak 383 data pada perhitungannya data dibagi kedalam data latih dan data uji sebesar 80-20. Data latih sebanyak 306 data dan data uji sebanyak 77 data. Perhitungan model ANN dengan aktivasi ReLU ini menghasilkan akurasi sebesar 0.961 dengan nilai AUC sebesar 0.99 yang berarti model ANN memiliki kinerja sangat baik untuk digunakan.

Keywords: ANN, Klasifikasi, Tiroid

Abstract

The thyroid gland, which has an important role in the body as a substance that affects metabolism, is included in the field of endocrinology. There are 22,368 cases of death in the United States caused by thyroid cancer. The death rate must continue to be suppressed so that the death rate caused by thyroid disorders can be reduced. One of the ways to reduce the mortality rate is by using technological developments. The data mining algorithm which is a method of learning data in recognizing complex patterns from big data and also a variety of decision making in the clinical world is the artificial neural network algorithm. In this study, the authors will classify thyroid patients who can relapse or not and find out what factors are associated with the recovery of thyroid disorder patients this study uses the Artificial neural network algorithm model with ReLU activation. This study uses thyroid which is secondary data from kaggle.com as much as 383 data in the calculation of the data is divided into training data and test data by 80-20. Training data is 306 data and test data is 77 data. The calculation of the ANN model with ReLU activation produces an accuracy of 0.961 with an AUC value of 0.99, which means that the ANN model has very good performance for use.

Keywords: ANN, Classification, Tiroid

1. Pendahuluan

Kelenjar tiroid yang memiliki peran penting dalam tubuh sebagai zat yang mempengaruhi metabolisme termasuk kedalam bidang endokrinologi (Ariawan et al., 2022). Penyakit dengan jenis

hipotiroidisme, hipertiroidisme dan nodul tiroid dapat kambuh kapan saja dan berpengaruh pada kesehatan pasien atau orang yang memiliki penyakit berhubungan dengan tiroid (Zhou et al., 2019).

Dalam hasil penelitian (Beynon & Pinneri, 2016) terdapat 22.368 kasus

kematian di Amerika Serikat yang diakibatkan kanker tiroid. Angka ini lebih kecil dibandingkan dengan penyebab kematian non-kanker pada penderita kanker tiroid tersebut. Tetapi, angka kematian tersebut harus terus ditekan agar tingkat kematian yang disebabkan oleh kelainan tiroid bisa berkurang. Satu diantara cara penekanan angka kematian tersebut bisa menggunakan perkembangan teknologi.

Kanker tiroid merupakan jenis kanker yang memiliki tingkat kesembuhan yang tinggi, tetapi pasien memiliki 20-30% mengalami kekambuhan pada kanker tiroid yang mempengaruhi kualitas hidup mereka dan memerlukan intervensi medis lebih lanjut. Memprediksi kekambuhan kanker tiroid ini merupakan tantangan yang cukup kompleks dikarenakan banyak pengaruh yang berkaitan dengan kekambuhan pasien seperti yang akan dibahas dalam penelitian ini yaitu usia pasien, gender pasien, merokok atau tidak, riwayat merokok pasien, dll.

Dalam penggunaan metode konvensional seperti pencatatan prediksi kekambuhan pasien tiroid dengan cara penilaian klinis dan penggambaran tindak pengobatan pasien akan memiliki keterbatasan dalam akurasi dan menyebabkan kegagalan dalam prediksinya. Dengan meningkatnya datamedis yang tersedia peneliti memiliki peluang dalam menggunakan metode teknik pembelajaran mesin atau machine learning dengan model tertentu agar menghasilkan tingkat akurasi yang baik.

Klasifikasi kelainan tiroid pernah diklasifikasikan oleh Monaco dalam (Monaco, 2003), dalam penelitiannya Monaco membuat tabel yang membedakan masalah tiroid, dengan hasil tiroid dibedakan kedalam dua kelompok kelas. Kelas kutub dan sementara. Klasifikasi tiroid menurut Monaco harus dilakukan lagi setelah 30 tahun kedepan agar dapat melihat perkembangan berdasarkan faktor-faktor yang mempengaruhi kelenjar tiroid. Berikutnya penelitian mengenai tiroid pernah dibuat juga oleh Ariawan Eka (Ariawan et al., 2022) berbentuk sistem pakar untuk mengetahui penderita kelainan tiroid dengan bantuan teknologi.

Satu diantara perkembangan teknologi dalam dunia medis adalah pemanfaatan AI (Hassan et al., 2022). Algoritma data mining yang menjadi metode pembelajaran sebuah data dalam mengenali

pola yang kompleks dari bigdata dan juga beragam dalam pengambilan keputusan dalam dunia klinis adalah algoritma artificial neural network (Wichmann et al., 2020).

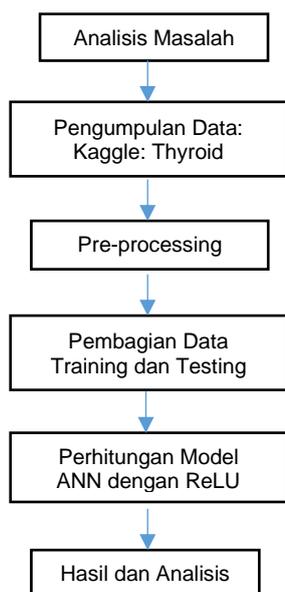
Pada penelitian ini akan mengklasifikasikan penderita tiroid yang bisa kambuh atau tidak. Dengan gap data yang cukup banyak sekitar 383 data pasien tiroid dibandingkan dengan penelitian sebelumnya dan juga menggunakan model Artificial Neural Network dengan 7 node inputan, dua hidden layer, dua layer output dan menggunakan aktivasi ReLU yang akan meningkatkan akurasi klasifikasi kekambuhan pasien tiroid.

Pada penelitian sebelumnya *artificial neural network* atau yang biasa dikenal dengan ANN digunakan para peneliti untuk menyelesaikan banyak masalah (Wang & Summers, 2012) Baik klasifikasi maupun juga prediksi. Artificial neural network juga dapat digunakan dengan berbagai jenis data, seperti data dengan tipe numeric dan image (VanderPlas, 2017). Seperti yang dilakukan oleh Heningtyas (Prabowo et al., 2021) dalam penelitian mengenai klasifikasi image tumbuhan obat Keji Beling yang menghasilkan tingkat akurasi sebesar 82.5%. ANN juga digunakan pada 2022 yang memberikan pengetahuan bahwa model ANN juga dapat digunakan untuk mengolah data berupa data text yang digunakan untuk membuat chatbot dengan pendekatan NLP. (Purwitasari & Soleh, 2022).

Maka dari itu pada penelitian ini, akan mengimplementasikan data tiroid seperti yang dilakukan Monaco tahun 2023, dan Ariawan et al tahun 2022, dengan tujuan dapat mendeksi lebih dini pada pasien yang beresiko mengalami gangguan tiroid. Pada penelitian ini juga dilakukan menggunakan 7 node dengan 2 hidden layer menggunakan model algoritma artificial neural network seperti penelitian Wichmann et al tahun 2020, Wang & Summers tahun 2012, VanderPlas tahun 2017, serta Purwitasari & Soleh pada tahun 2022, dengan tujuan penelitian untuk memberikan dasar dalam pengambilan keputusan.

2. Metode Penelitian

Proses penelitian ini digambarkan oleh alur penelitian pada gambar 1 seperti berikut:



Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 1. Alur Penelitian

Gambar 1 menjelaskan alur penelitian yang diawali oleh:

Analisis Masalah:

Tingkat kematian yang cukup tinggi dikarenakan ketidaktahuan pasien akan penyakitnya membutuhkan bantuan teknologi dalam bentuk klasifikasi dari big data history pasien-pasien yang mengalami kelainan tiroid. maka perlu dibuat penelitian mengenai klasifikasi pasien yang bisa kambuh atau tidak menggunakan model algoritma Artificial neural network agar nilai akurasi lebih tinggi dan dapat diimplementasikan.

Pengumpulan Data:

Data diperoleh dari kaggle.com dengan jumlah data sebanyak 383 data pasien yang mengalami kelainan tiroid.

Pre-Processing:

Pada tahap ini data dilakukan pemeriksaan missing value dari data yang tersedia. Setelah dirasa cukup dan tidak ada missing value, berikutnya merubah data atau proses encoding data. Proses ini dilakukan agar perhitungan bisa dilakukan dengan numerical data.

Perhitungan Artificial Neural Network:

Pada tahap ini data diolah menggunakan model ANN. Data latih dilatih menggunakan fungsi MCLClassifier. Dengan menggunakan aktivasi ReLU.

Dari perhitungan ini kita akan melihat implementasi data categorical yang dihitung dirubah kedalam data numerical sehingga bisa dilakukan perhitungan ANN dan melihat peningkatan akurasinya.

3. Hasil dan Pembahasan

3.1. Pengolahan Data

Data mining merupakan proses yang mempelajari data dari data yang telah dilatih. Pada penelitian ini pelatihan dan pengujian data dibagi kedalam 80% data pelatihan dan 20% data uji. Dimana 80% dari 383 data sebanyak 306 data untuk pelatihan. 20% dari 383 sebanyak 77 data untuk data uji.

Tabel 1. Dataset Tiroid

Age	Gender	Smoking	Hx Smoking	Hx Radiotherapy	Thyroid Function	Response	Recurred
27	F	No	No	No	Euthyroid	Indeterminate	No
34	F	No	Yes	No	Euthyroid	Excellent	No
30	F	No	No	No	Euthyroid	Excellent	No
62	F	No	No	No	Euthyroid	Excellent	No
62	F	No	No	No	Euthyroid	Excellent	No
52	M	Yes	No	No	Euthyroid	Indeterminate	No
....
....
72	M	Yes	Yes	Yes	Euthyroid	Biochemical Incomplete	Yes

81	M	Yes	No	Yes	Euthyroid	Structural Incomplete	Yes
72	M	Yes	Yes	No	Euthyroid	Structural Incomplete	Yes
61	M	Yes	Yes	Yes	Clinical Hyperthyroidism	Structural Incomplete	Yes
67	M	Yes	No	No	Euthyroid	Structural Incomplete	Yes

Sumber: <https://www.kaggle.com/datasets/jainaru/thyroid-disease-data>

Tabel 1 merupakan data thyroid sebanyak 383 data. Langkah berikutnya adalah pre-processing data.

3.2. Pre-processing Data

Pre-processing data yang dilakukan untuk data diatas adalah menghitung missing value.

```
missing_values_count=df.isnull().sum()
print(missing_values_count)
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 2. Perintah pre-processing missing value

Pada Gambar 2 merupakan perintah untuk melihat apakah ada missing value pada data. Data ini sudah sangat baik, jadi setiap data terisi dengan baik dan tidak memiliki missing value. Berikutnya adalah encoding. Encoding atau merubah data dari categorical menjadi format yang dapat digunakan oleh algoritma. Perintah tersebut dijelaskan seperti ini:

```
[11] ohe_genders = pd.get_dummies(df[['Gender']])

# Apply to numeric to each column individually
for col in ohe_genders.columns:
    ohe_genders[col] = pd.to_numeric(ohe_genders[col])

print(ohe_genders)
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 3. Perintah one hand-encoding

Pada gambar 3 adalah contoh one hand-encoding untuk field gender. Data yang tadinya berbentuk F dan M diubah kedalam bilangan binner. Berikut hasilnya:

```
Gender_F  Gender_M
0         True   False
1         True   False
2         True   False
3         True   False
4         True   False
..        ...     ...
378       False  True
379       False  True
380       False  True
381       False  True
382       False  True

[383 rows x 2 columns]
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 4. Hasil perhitungan one hand-encoding field gender

One hand-encoding terus dilanjutkan dengan perintah yang sama untuk field Smoking, Hx Smoking, dan juga Hx Radiotherapy. One hand-encoding juga bisa digunakan untuk data dengan record lebih dari tiga jenis seperti pada field Thyroid Functin dan juga Response.

```
Thyroid Function_Clinical Hyperthyroidism \
0         False
1         False
2         False
3         False
4         False
..        ...
378       False
379       False
380       False
381         True
382         False

Thyroid Function_Clinical Hypothyroidism  Thyroid Function_euthyroid \
0         False                          True
1         False                          True
2         False                          True
3         False                          True
4         False                          True
..        ...                            ...
378       False                          True
379       False                          True
380       False                          True
381         False                         False
382         False                         True

Thyroid Function_Subclinical Hyperthyroidism \
0         False
1         False
2         False
3         False
4         False
..        ...
378       False
379       False
380       False
381         False
382         False

Thyroid Function_Subclinical hypothyroidism
0         False
1         False
2         False
3         False
4         False
..        ...
378       False
379       False
380       False
381         False
382         False

[383 rows x 5 columns]
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 5. Hasil perhitungan one hand-encoding field Thyroid Function

Proses encoding ini menjabarkan data menjadi bilangan binner dengan mengganti nilai dari field tersebut menjadi false(0) jika salah dan true(1) jika benar.

Setelah didapatkan hasil encoding, penelitian pre-processing masih dilanjutkan menggunakan langkah MinMaxScaler. perintah ini sekaligus membagi data latih dan data uji menjadi 80%-20%.

```
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
x_train = MinMaxScaler().fit_transform(x_train)
x_test = MinMaxScaler().fit_transform(x_test)
print(x_train)
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 6. Perintah MinMaxScaler dan Pembagian data latih dan uji.

MinMaxScaler mengubah angka menjadi antara 0 sampai 1 agar data tidak bias. Data setelah MinMaxScaler menjadi seperti ini:

```
[[0.35384615 1. 0. ... 1. 0. 0. ]
 [0.69238769 1. 0. ... 0. 0. 1. ]
 [0.18461538 1. 0. ... 0. 1. 0. ]
 ...
 [0.33846154 1. 0. ... 0. 1. 0. ]
 [0.63076923 1. 0. ... 0. 0. 1. ]
 [0.15384615 1. 0. ... 1. 0. 0. ]]
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 7. Hasil perhitungan minmaxscaller

3.3. Pemodelan ANN

Proses selanjutnya adalah model ANN. Pada penelitian ini peneliti menggunakan Multi-Layer Perceptron Classifier (MCLClassifier) untuk melatih model dalam mempelajari pola dalam data sebagai library yang ada pada sklearn. Perintahnya digambarkan pada gambar 8.

```
clf = MCLClassifier(activation='relu', solver='sgd', hidden_layer_sizes=2, learning_rate_init=0.5, max_iter=100)
clf.fit(x_train, y_train)
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 8. Model ANN

Pada gambar 8 perintah pemodelan ANN diberikan. Pada perintah tersebut menggambarkan penggunaan MCLClassifier, fungsi ini digunakan untuk memudahkan bagian input layer secara otomatis. Aktivasi yang digunakan pada penelitian ini adalah ReLU(Rectified Linear Unit), aktivasi ini akan menjadikan nilai output sama dengan nilai input jika nilai inputnya bernilai positif, tetapi nilai output akan dijadikan 0(no) jika nilai inputnya bernilai negatif. Selanjutnya untuk solver

pada penelitian ini menggunakan SGD agar nilai akurasi bisa lebih optimal. Berikutnya penentuan hidden layer. Hidden layer untuk penelitian ini dibuat kedalam dua hidden layer dengan learning ratenya sebesar 0.5. lalu untuk uji akurasi data ini akan diuji sebanyak 100 kali sesuai nilai max-iter yang diisikan.

Hasil dari pelatihan diatas ditunjukkan oleh gambar berikut:

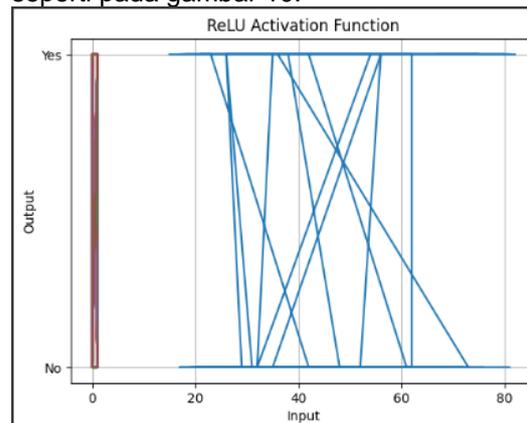
```
y_pred=clf.predict(x_test)
print(y_pred)
```

```
['No' 'No' 'Yes' 'Yes' 'No' 'Yes' 'No' 'No' 'Yes' 'Yes' 'No' 'No' 'No'
 'No' 'No' 'No' 'No' 'Yes' 'Yes' 'No' 'No' 'No' 'No' 'No' 'No' 'No'
 'No' 'No' 'Yes' 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'Yes' 'Yes' 'No'
 'No' 'No' 'No' 'No' 'No' 'No' 'Yes' 'No' 'No' 'No' 'Yes' 'No'
 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'Yes' 'No' 'Yes' 'No' 'Yes'
 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'No' 'Yes' 'No']
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 9. Hasil prediksi data latih

Sesuai dengan teori mengenai ReLU nilai prediksi diatas jika digambarkan menggunakan diagram akan berbentuk seperti pada gambar 10:



Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 10. Diagram aktivasi ReLU

Selanjutnya data uji dengan menghitung akurasi:

```
round(accuracy_score(y_test,y_pred),3)
```

```
0.961
```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 11. Perintah dan hasil perhitungan Akurasi

Akurasi untuk pengujian data uji sebesar 0.961. dengan nilai akurasi tersebut

menandakan bahwa perhitungan sangat baik. Selebihnya kita hitung nilai recall dan precisionnya dengan perintah berikut:

```

y_pred = clf.predict(x_test)
accuracy = accuracy_score(y_test, y_pred)

print("Accuracy:", accuracy)
print("\nClassification Report:\n", classification_report(y_test, y_pred))
print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))

Accuracy: 0.961038961038961

Classification Report:
              precision    recall  f1-score   support
   No         0.95         1.00         0.97         58
   Yes         1.00         0.04         0.91         19

 accuracy
macro avg   0.98         0.92         0.94         77
weighted avg 0.96         0.96         0.96         77

Confusion Matrix:
[[ 58  0]
 [  3 16]]

```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 12. Perhitungan Recall dan Precision

Setelah pengujian model nilai recall dan precision juga dihitung untuk evaluasi model ANN terhadap data uji. Dilanjutkan dengan perhitungan ROC dan AUC untuk melihat apakah nilai ROC. Berikut perintahnya:

```

y_test_numeric = [1 if label == 'Yes' else 0 for label in y_test]
y_probs = clf.predict_proba(x_test)[:, 1]
fpr, tpr, thresholds = roc_curve(y_test_numeric, y_probs)
auc = roc_auc_score(y_test_numeric, y_probs)
print("AUC", auc)

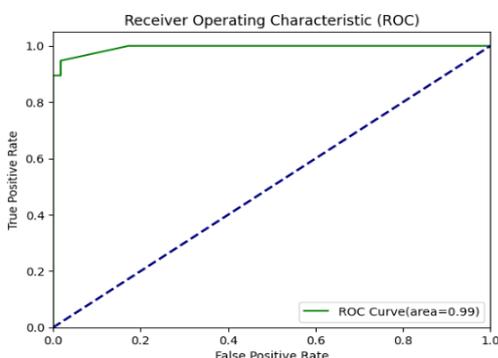
AUC 0.9941016333938294

```

Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 13. Perintah perhitungan AUC

Perhitungan pada gambar 13 menjelaskan bahwa model dengan data tiroid ini mendapatkan nilai AUC sebesar 0.99 mendekati 1, itu berarti pemodelan ini memiliki kinerja yang sangat baik. Berikutnya kita hitung ROC untuk penelitian ini. Berikut kurvanya:



Sumber: (Yusuf & Hidayatulloh, 2024)

Gambar 14. Kurva ROC

4. Kesimpulan

Klasifikasi kelainan tiroid dapat kambuh atau tidak, dapat diuji dengan baik. Sehingga masalah dengan nilai uji yang baik nantinya akan bisa memberikan info kepada pasien pengidap tiroid, akan lebih mengetahui apakah dengan faktor-faktor pembawa tersebut tiroidnya akan kambuh atau tidak. Sehingga pasien tiroid akan lebih memperhatikan faktor kesehatan mereka, yang akan menahan kambuhnya tiroid yang mereka miliki. Dengan model Artificial Neural Network (ANN) menggunakan 7 node inputan, dua hidden layer, dua layer output dan menggunakan aktivasi ReLU. Dengan hasil akurasi sebesar 0.961 dan nilai AUC sebesar 0.99 memberikan kesimpulan bahwa model ANN ini sangat dapat digunakan untuk jenis data sekunder dan dapat diimplementasikan kedalam program sederhana. Penelitian selanjutnya dengan data ini peneliti bisa melakukan klasifikasi kategori resiko kanker.

5. Referensi

- Ariawan, W. E., Made, I., & Putra, A. W. (2022b). Sistem Pakar Mendiagnosa Penyakit Tiroid Menggunakan Metode Certainty Factor Berbasis Web. *Jurnal Sutasoma*. <https://s.id/journalsutasoma>
- Beynon, M. E., & Pinneri, K. (2016). An Overview of the Thyroid Gland and Thyroid-Related Deaths for the Forensic Pathologist. In *Academic Forensic Pathology* (Vol. 6, Issue 2, pp. 217–236). SAGE Publications Inc. <https://doi.org/10.23907/2016.024>
- Hassan, C., Balsamo, G., Lorenzetti, R., Zullo, A., & Antonelli, G. (2022). Artificial Intelligence Allows Leaving-In-Situ Colorectal Polyps. *Clinical Gastroenterology and Hepatology*, 20(11), 2505-2513.e4. <https://doi.org/10.1016/j.cgh.2022.04.045>
- Monaco, F. (2003). Clinical perspective: Classification of thyroid diseases: Suggestions for a revision. In *Journal of Clinical Endocrinology and Metabolism* (Vol. 88, Issue 4, pp. 1428–1432). <https://doi.org/10.1210/jc.2002-021260>
- Prabowo, R., Heningtyas, Y., Yusman, M., Iqbal, M., Dwi, O., Wulansari, E., Ilmu, J., Lampung, U., & Matematika, J. (2021). Klasifikasi Image Tumbuhan Obat (Keji Beling) Menggunakan

-
- Artificial Neural Network. In *Jurnal Komputasi* (Vol. 9, Issue 2).
- Purwitasari, N. A., & Soleh, M. (2022). *Implementasi Algoritma Artificial Neural Network Dalam Pembuatan Chatbot Menggunakan Pendekatan Natural Language Processing (Implementation Of Artificial Neural Network Algorithm In Chatbot Development Using Natural Language Processing Approach)*.
- VanderPlas, J. (2017). *Python Data Science Handbook*. www.allitebooks.com
- Wang, S., & Summers, R. M. (2012). Machine learning and radiology. In *Medical Image Analysis* (Vol. 16, Issue 5, pp. 933–951). <https://doi.org/10.1016/j.media.2012.02.005>
- Wichmann, J. L., Willemink, M. J., & De Cecco, C. N. (2020). Artificial Intelligence and Machine Learning in Radiology: Current State and Considerations for Routine Clinical Implementation. In *Investigative Radiology* (Vol. 55, Issue 9, pp. 619–627). Lippincott Williams and Wilkins. <https://doi.org/10.1097/RLI.00000000000000673>
- Yusuf, L., & Hidayatulloh, T. (2024). *Laporan Penelitian: Implementasi Algoritma Artificial Neural Network dengan Aktivasi ReLU (Klasifikasi Tiroid)*.
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2019). Unet++: A nested u-net architecture for medical image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11045 LNCS, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1