

Perbandingan Metode *Artificial Neural Network* dan *Convolutional Neural Network* untuk Memprediksi Jumlah Distribusi Air di PDAM Kota Denpasar

I Putu Agung Bayupati^{1*}, Anak Agung Ayu Sintya Dewi^{2*}, Ni Kadek Ayu Wirdiani³ 

^{1,2,3} Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana, Badung, Bali, Indonesia

ARTICLE INFO

Article history:

Received June 06, 2022

Revised June 11, 2022

Accepted April 23, 2023

Available online July 25, 2023

Kata Kunci:

Artificial Neural Network,
Convolutional Neural Network,
Distribusi Air

Keywords:

Artificial Neural Network,
Convolutional Neural Network,
Water Distributio



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 20223 by Author. Published by Universitas Pendidikan Ganesha.

ABSTRAK

Kelangsungan hidup di dunia pasti tidak akan bisa terlepas dari penggunaan air. Seiring dengan berkembangnya zaman tentu akan diikuti dengan bertambahnya jumlah penduduk yang juga akan berdampak terhadap meningkatnya kebutuhan air bersih, salah satunya di Kota Denpasar. PDAM merupakan salah satu instansi yang melayani ketersediaan air bersih. Tujuan diadakannya penelitian ini untuk melakukan peramalan terhadap jumlah distribusi air di PDAM Kota Denpasar dengan membandingkan metode ANN (*Artificial Neural Network*) serta CNN (*Convolutional Neural Network*) yang juga melibatkan proses seleksi fitur menggunakan metode SFS (*Sequential Forward Selection*). Pendekatan kuantitatif digunakan untuk melakukan proses peramalan terhadap jumlah distribusi air (Y) dengan melibatkan fitur jumlah produksi (X_1), kebocoran (X_2), pembelian (X_3), dan pelanggan air (x_4). Hasil seleksi fitur dengan metode SFS menunjukkan bahwa jumlah produksi air (X_1) dan kebocoran air (X_2) dengan tingkat kesalahan sebesar 0.031069 tepat digunakan untuk membentuk model peramalan jumlah distribusi air dengan metode ANN yang menghasilkan nilai MSE dan MAPE berturut-turut sebesar 0.040 serta 2.02%. Berdasarkan hasil yang diperoleh, model yang dikembangkan termasuk model peramalan yang sangat baik, sehingga tepat digunakan untuk melakukan proses peramalan jumlah distribusi air.

ABSTRACT

One of the most essential part in human's life sustainability is the existence of water. The population of human is gradually increasing time to time that it will affect the demand of clean water, as one of the case happens in Denpasar city. PDAM (*Perusahaan Daerah Air Minum*) is one of the institution that provides the distribution and accessibility for clean water. The purpose of this study is to forecast the amount of water distribution in PDAM Denpasar City by comparing the ANN (*Artificial Neural Network*) method with CNN (*Convolutional Neural Network*) which also involves a feature selection process using the SFS (*Sequential Forward Selection*) method. A quantitative approach is used to forecast the amount of water distribution (Y) by involving the features of the amount of water production (X_1), water leakage (X_2), water purchases (X_3), and water customers (x_4). The results of feature selection using the SFS method show that the amount of water production (X_1) and water leakage (X_2) with an error rate of 0.031069 is accurate to form a forecasting model for the amount of water distribution using the ANN method which produces MSE and MAPE values of 0.040 and 2.02% respectively. Based on the results obtained, the developed model includes a very good forecasting model, so it is appropriated to use the process of forecasting the amount of water distribution.

1. PENDAHULUAN

Air merupakan kebutuhan dasar bagi kelangsungan hidup manusia. Air bersih digunakan oleh penduduk untuk menunjang berbagai kegiatan sehari-hari dalam lingkungan masyarakat (Kurniawan et al., 2021; Lestari et al., 2021). Salah satu cara untuk mendapatkan air bersih adalah melalui layanan yang disediakan oleh pemerintah, yaitu PDAM (*Perusahaan Daerah Air Minum*). Jumlah distribusi air di PDAM Kota Denpasar pada Desember 2020 telah mencapai 3.139.097 liter dengan jumlah pelanggan sebanyak 88.820 unit. Terhitung pada tahun 2020 penduduk di Kota Denpasar telah mencapai 725.314 jiwa berdasarkan data yang diperoleh dari Badan Pusat Statistik Provinsi Bali. Seiring dengan berjalannya waktu, tidak menutup kemungkinan terjadinya peningkatan

*Corresponding author.

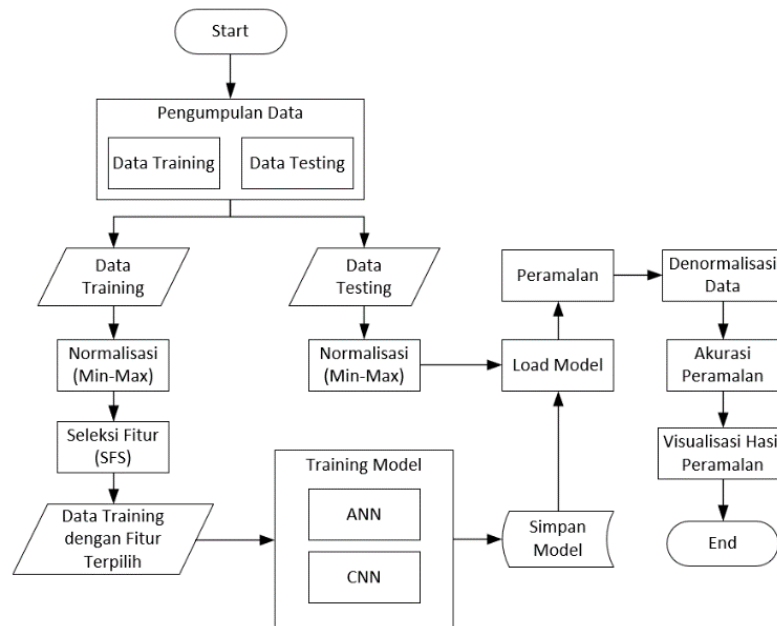
E-mail addresses: ayusintyaadewi@gmail.com (I Putu Agung Bayupati)

jumlah penduduk yang tentunya akan berdampak terhadap meningkatnya jumlah pelanggan di PDAM Kota Denpasar. Oleh karena itu, untuk mengetahui jumlah kebutuhan air bersih di masa mendatang maka dapat dilakukan proses peramalan. Seiring dengan meningkatnya jumlah pelanggan, tentu pihak PDAM harus mampu untuk memenuhi kebutuhan air bersih yang dibutuhkan oleh para pelanggannya. Oleh sebab itu, diperlukan proses peramalan mengenai jumlah distribusi air yang akan disalurkan ke depannya. Proses peramalan ini bertujuan untuk memberikan gambaran kepada pihak PDAM guna mengambil keputusan di masa mendatang berdasarkan hasil peramalan yang diperoleh. Proses peramalan dilakukan dengan melibatkan data historis yang sudah ada dari tahun sebelumnya (Rachman, 2018; Sagheer & Kotb, 2019). Salah satu metode peramalan yang kini populer digunakan adalah neural network seperti ANN (*Artificial Neural Network*) dan CNN (*Convolutional Neural Network*). *Neural network* bekerja seperti jaringan saraf manusia dan mampu mengolah informasi untuk menghasilkan output yang diharapkan (Abiodun et al., 2018; Xu et al., 2020). Berdasarkan penelitian terdahulu mengenai peramalan jumlah distribusi air menggunakan metode ANFIS (*Adaptive Neuro Fuzzy Inference System*) (Oktafiansyah & Uperiati, 2021) masih terdapat kekurangan, di mana variabel data yang diteliti masih bersifat individu tanpa melibatkan variabel terkait lainnya. Oleh karena itu, dalam penelitian ini dikembangkan kembali model analisis dengan melibatkan variabel independen (X) untuk memprediksi variabel dependen (Y). Variabel independen (X) dipilih melalui proses seleksi fitur dengan metode SFS (*Sequential Forward Selection*). Variabel terpilih kemudian diolah dengan metode ANN dan CNN yang nantinya hasil yang diperoleh akan dibandingkan untuk menentukan metode peramalan mana yang terbaik digunakan untuk memprediksi jumlah distribusi air di masa mendatang. Kombinasi metode seleksi fitur SFS dengan neural network ini diharapkan mampu memaksimalkan kinerja dari model peramalan yang dibentuk, sehingga dapat menghasilkan nilai peramalan yang akurat. Metode SFS telah banyak digunakan untuk menentukan kombinasi antarvariabel yang memiliki hubungan kuat untuk meminimalisir dilibatkannya variabel independen yang minim kaitannya dengan variabel dependen (Martín-Vázquez et al., 2018; Shafiee et al., 2021). Beragam masalah peramalan di berbagai bidang juga sudah banyak diselesaikan dengan melibatkan metode ANN (Attoue et al., 2018; Ozoegwu, 2019; Toğa et al., 2021) serta CNN (Huang et al., 2020; Pan & Zhou, 2020; Yazici et al., 2022). Keuntungan dari metode ANN yaitu dapat mengolah data linier maupun non linier, memiliki kemampuan untuk mengenali interaksi yang terjadi antara variabel yang diteliti, serta mampu memecahkan masalah yang sulit dilakukan dengan menggunakan metode peramalan tradisional (Bhakti, 2019; Trach et al., 2021). Metode peramalan CNN sendiri memiliki keuntungan untuk mengekstrak fitur secara otomatis selama proses pembentukan model sehingga dapat mempersingkat waktu selama proses berlangsung (Pan & Zhou, 2020; Zhao et al., 2019). Hasil peramalan dengan ANN mampu menghasilkan model yang dapat memecahkan masalah dengan nilai MPE (*Mean Percentage Error*) sebesar 5.67% (Ozoegwu, 2019). Untuk pemecahan masalah peramalan dengan metode CNN mampu menghasilkan nilai MAPE (*Mean Absolute Percentage Error*) sebesar 2.21% (Yazici et al., 2022). Berdasarkan kajian tersebut, maka kemudian metode SFS, ANN, serta CNN dapat diterapkan pada kasus peramalan jumlah distribusi air. Implementasi kombinasi metode yang digunakan pada penelitian ini menjadi keterbaruan yang dapat memberikan perbedaan dari kajian sebelumnya (Oktafiansyah & Uperiati, 2021) yang hanya memaparkan kinerja dari satu metode yang masih sederhana tanpa melibatkan keberadaan variabel independen. Selain itu, akurasi hasil peramalan yang dicantumkan kajian terdahulu hanya berupa nilai RMSE sebesar 0.07991 yang tentunya kurang bisa dipahami oleh khalayak umum. Dengan minimnya kajian mengenai peramalan jumlah distribusi air, tentu diperlukan kembali pembaharuan mengenai kajian terkait sebagai pembanding dengan mengembangkan model analisis dari penelitian yang sebelumnya. Tujuan dilakukan penelitian ini yaitu untuk memprediksi jumlah distribusi air di PDAM Kota Denpasar dengan membandingkan metode peramalan ANN dengan CNN berdasarkan metode seleksi fitur SFS. Nilai peramalan yang diperoleh diharapkan mampu memberikan gambaran bagi pihak PDAM Kota Denpasar dalam mempertimbangkan jumlah air yang akan didistribusikan kepada para pelanggan di masa

2. METODE

Penelitian ini menerapkan pendekatan kuantitatif yang melibatkan perhitungan matematis. Data penelitian didapat dari PDAM Kota Denpasar yang bersifat bulanan sejak Bulan Januari 2015 hingga Desember 2020 berjumlah 72 baris data. Variabel data penelitian yang diperoleh yaitu jumlah produksi air (X1), kebocoran air (X2), pembelian air (X3), pelanggan air (X4), dan distribusi air (Y). Satuan yang digunakan untuk jumlah distribusi, produksi, kebocoran, dan pembelian air adalah kubik (m^3). Sementara

itu untuk jumlah pelanggan memiliki satuan unit. Proses peramalan jumlah distribusi air pada penelitian ini digambarkan melalui model analisis yang dapat dilihat pada Gambar 1.



Gambar 1. Model Analisis Peramalan Jumlah Distribusi Air

Gambar 1 merupakan model analisis dalam melakukan proses peramalan jumlah distribusi air yang tentunya melalui berbagai tahapan. Proses dimulai dari melakukan pengumpulan data yang kemudian diperoleh lima variabel penelitian, yaitu jumlah distribusi, produksi, kebocoran, pembelian, dan pelanggan air. Data penelitian yang diperoleh kemudian dibagi menjadi data *training* dan data *testing*. Data *training* yang digunakan yaitu dari Bulan Januari 2015 hingga Desember 2019 sebanyak 60 baris data. Data *testing* yang digunakan yaitu dari Bulan Januari 2020 hingga Desember 2020 sebanyak 12 baris data. Proses yang dilakukan selanjutnya adalah melakukan normalisasi data menggunakan metode Min-Max Normalization. Data *training* kemudian melalui tahap seleksi fitur menggunakan metode SFS untuk selanjutnya melalui tahap *training* model dengan pengujian menggunakan metode ANN serta CNN. Model yang didapat kemudian digunakan untuk melakukan proses peramalan jumlah distribusi air. Hasil peramalan yang diperoleh kemudian didenormalisasi untuk mengembalikan nilai jumlah distribusi air ke dalam rentang semula. Proses selanjutnya yang dilakukan adalah mengukur tingkat akurasi model peramalan dengan MSE dan MAPE guna mengetahui model peramalan jumlah distribusi air terbaik dan akurat. Proses terakhir yang dilakukan adalah visualisasi data hasil peramalan berdasarkan metode peramalan jumlah distribusi air terbaik.

Min-Max Normalization

Proses normalisasi data dilakukan dengan menggunakan metode Min-Max. Proses normalisasi data ditujukan untuk mengubah rentang nilai setiap variabel data penelitian agar berada pada skala atau rentang nilai 0 hingga 1. Persamaan dari Min-Max Normalization dapat dilihat pada Persamaan 1.

$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{range.max} - \text{range.min}) + \text{range.min} \tag{1}$$

Persamaan 1 merupakan formula yang digunakan untuk mengubah rentang nilai aktual menjadi skala 0 hingga 1. Variabel v' merupakan nilai setelah dilakukannya proses normalisasi. Proses kalkulasinya yaitu nilai aktual (v) dikurangi dengan nilai terendah (\min_A) pada variabel penelitian yang dimaksud lalu dibagi dengan selisih nilai antara nilai tertinggi (\max_A) dengan nilai terendah (\min_A) pada variabel penelitian yang dimaksud. Hasil tersebut kemudian dikalikan dengan selisih skala tertinggi yaitu bernilai 1 (range.max) dikurangi skala terendah yaitu bernilai 0 (range.min). Proses terakhir yaitu hasil tersebut kemudian dijumlahkan dengan nilai 0 (range.min) (Altan et al., 2019; Zhu et al., 2020). Hasil dari proses peramalan yang masih berada pada rentang nilai 0 hingga 1, nantinya akan dikembalikan ke dalam rentang nilai semula. Proses mengubah nilai ke dalam rentang semula dikenal dengan proses denormalisasi data. Proses denormalisasi data dilakukan dengan mengikuti formula pada Persamaan 2.

$$v = v' * (max_v - min_v) + min_v \tag{2}$$

Persamaan 2 merupakan formula yang digunakan untuk mengubah rentang nilai normalisasi 0 hingga 1 kembali menjadi ke rentang nilai semula. Proses perhitungan yang dilakukan yaitu perkalian antara nilai hasil peramalan dengan nilai selisih nilai aktual tertinggi (max_v) dikurangi nilai aktual terendah (min_v). Hasil tersebut kemudian dikurangi dengan nilai data aktual terendah (min_v) untuk mengembalikan nilai hasil peramalan ke dalam rentang semula (Priliani et al., 2018).

SFS (Sequential Forward Selection)

Metode seleksi fitur yang digunakan adalah SFS (*Sequential Forward Selection*). Kombinasi fitur dengan tingkat kesalahan terendah nantinya digunakan sebagai prediktor pada proses *training* model. Algoritma dari metode SFS dapat dilihat pada Tabel 1.

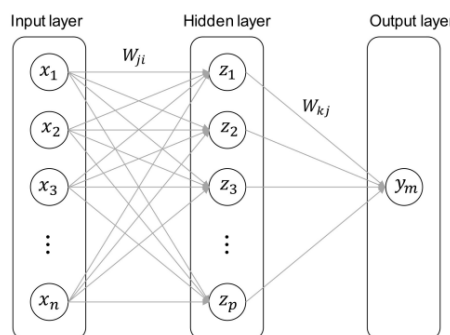
Tabel 1. Algoritma Sequential Forward Selection

Algoritma SFS (Sequential Forward Selection)	
1	$Y \leftarrow \{ \}; F \leftarrow features$
2	$while\ size(F) > 0$
3	$f = argmin_{f \in F} [J(Y + f)]$
4	$Y \leftarrow Y + f$
5	$F \leftarrow F - f$

Tabel 1 merupakan algoritma dari metode seleksi fitur SFS (*Sequential Forward Selection*). Pemodelan dimulai dari nol fitur yang kemudian secara bertahap ditambahkan satu fitur sampai jumlah fitur terakhir terpenuhi. Setiap iterasi yang terjadi merupakan lanjutan dari kombinasi fitur sebelumnya dengan tingkat kesalahan terendah (Castangia et al., 2021).

ANN (Artificial Neural Network)

ANN (*Artificial Neural Network*) atau JST (*Jaringan Saraf Tiruan*) merupakan sistem yang serupa dengan jaringan saraf manusia untuk mengolah informasi agar menghasilkan *output* yang diinginkan (Freire et al., 2019). ANN merupakan model nonlinier sehingga dapat menghasilkan solusi yang lebih realistis untuk membantu mengatasi masalah di kehidupan nyata (Güven & Şimşir, 2020). ANN termasuk ke dalam bagian dari *artificial intelligence*, yang di mana topologi paling umum dari ANN yaitu MLP (*Multilayer Perceptron*). MLP disusun atas beberapa *layer* diantaranya yaitu *input layer*, *hidden layer*, dan *output layer*. Pada setiap *layer* tersusun atas *nodes* yang saling dihubungkan oleh *weight*.



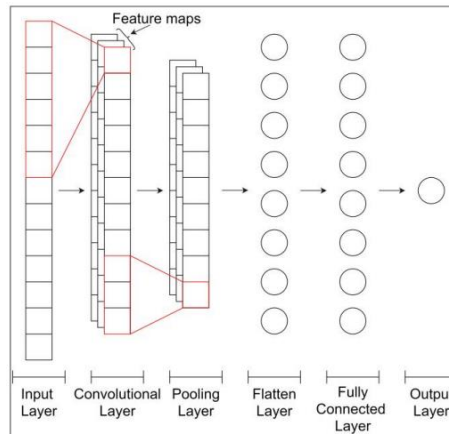
Gambar 2. Arsitektur Umum ANN (Multilayer Perceptron)

Gambar 2 merupakan gambaran dari arsitektur umum metode ANN (*Artificial Neural Network*) yang terdiri dari atas tiga *layer*. *Input layer* merupakan lapisan yang digunakan sebagai masukan informasi pada jaringan. Setiap *layer* dihubungkan oleh *weight* atau himpunan bobot. *Weight* diperoleh selama proses *training* model yang nilainya selalu berubah. *Hidden layer* ini bisa berjumlah lebih dari satu. Setelah melalui *hidden layer*, informasi kemudian diteruskan menuju *output layer* sebagai tempat keluaran informasi (Suryadibrata & Chandra, 2020).

CNN (Convolutional Neural Network)

CNN (*Convolutional Neural Network*) merupakan perkembangan lebih lanjut dari MLP (*Multilayer Perceptron*). Secara umum terdapat dua buah *layer* pada CNN, yaitu *feature extraction layer* dan *fully*

connected layer (Agga et al., 2021). Feature extraction layer terletak setelah input layer. Feature extraction layer disusun oleh dua jenis layer, yaitu convolutional layer dan pooling layer (Yan et al., 2021). Dalam berbagai penelitian, metode CNN banyak digunakan untuk mengolah data berupa citra atau gambar menggunakan 2D convolutional layer. Di samping itu, metode CNN juga bisa digunakan untuk mengolah data berupa vektor seperti sinyal serta data time series menggunakan 1D convolutional layer (Syarifudin et al., 2021).



Gambar 3. Arsitektur Umum CNN 1D

Gambar 3 merupakan gambaran dari arsitektur umum CNN (Convolutional Neural Network) 1D yang terdiri dari beberapa layer. Input layer merupakan lapisan yang digunakan sebagai masukan dari informasi yang akan diolah. Pada convolutional layer diterapkan filter pada fitur yang dimasukkan untuk membentuk feature map. Pooling layer kemudian digunakan untuk mengurangi dimensi dari feature map yang dihasilkan. Flatten layer digunakan untuk mengubah feature map menjadi 1D vektor yang kemudian diteruskan pada fully connected layer. Fully connected layer menggunakan MLP untuk mengolah data sehingga didapat hasil yang diinginkan.

MSE (Mean Squared Error)

MSE (Mean Squared Error) merupakan salah satu metode yang dapat digunakan untuk mengukur tingkat kesalahan model peramalan yang dibentuk. Cara kerja MSE yaitu dengan mencari nilai rata-rata kuadrat kesalahan dari peramalan yang dilakukan (Prayudani et al., 2019). Formula yang digunakan untuk menghitung nilai MSE dapat dilihat pada Persamaan 3.

$$MSE = \frac{\sum |a-b|^2}{n} \quad (3)$$

Persamaan 3 merupakan formula yang digunakan untuk menghitung nilai MSE. Nilai MSE didapat dengan melakukan kalkulasi yaitu total jumlah selisih absolut nilai aktual (a) dikurangi dengan nilai peramalan (b) yang dikuadratkan. Hasil yang didapat kemudian dibagi dengan jumlah data (n) dalam satu periode peramalan.

MAPE (Mean Absolute Percentage Error)

MAPE (Mean Absolute Percentage Error) juga merupakan salah satu metode yang dapat digunakan untuk mengukur tingkat kesalahan model peramalan yang dibentuk dalam bentuk yang lebih mudah dipahami, yaitu persentase. Cara kerja MAPE yaitu dengan mencari nilai rata-rata persentase kesalahan mutlak (Prayudani et al., 2019). Formula yang digunakan untuk menghitung nilai MAPE dapat dilihat pada Persamaan 4.

$$MAPE = \frac{\sum \frac{|a-b|}{a}}{n} \times 100\% \quad (4)$$

Persamaan 4 merupakan formula yang digunakan untuk menghitung nilai MAPE. Nilai MAPE didapat dengan melakukan kalkulasi yaitu total jumlah selisih absolut nilai data aktual (a) dikurangi dengan nilai peramalan (b) lalu dibagi dengan nilai data aktual (a). Hasil tersebut kemudian dibagi dengan jumlah data (n) dalam satu periode peramalan. Langkah terakhir, hasil kalkulasi tersebut kemudian dikali

dengan nilai 100%. Nilai MAPE yang diperoleh dapat dikategorikan sesuai dengan yang disajikan pada Tabel 2.

Tabel 2. Kriteria Nilai MAPE

Retang Nilai MAPE	Kriteria
< 10%	Model peramalan sangat baik
10% - 20%	Model peramalan baik
20% - 50%	Model peramalan wajar
> 50%	Model peramalan buruk dan tidak akurat

Tabel 2 menunjukkan rentang nilai MAPE dan kriterianya yang digunakan untuk mengevaluasi model peramalan. Nilai MAPE yang kurang dari 10% termasuk ke dalam model peramalan yang sangat baik. Nilai MAPE sebesar 10%-20% termasuk ke dalam model peramalan yang baik. Nilai MAPE sebesar 20%-50% termasuk ke dalam model peramalan yang wajar. Nilai MAPE yang lebih besar dari 50% termasuk ke dalam model peramalan yang buruk dan tidak akurat (La Foucade et al., 2019).

3. HASIL DAN PEMBAHASAN

Hasil

Proses peramalan jumlah distribusi air dimulai dengan tahap pembagian *dataset* menjadi data *training* dan data *testing*. Proses dilanjutkan dengan melakukan normalisasi data *training* dan data *testing* menggunakan metode Min-Max Normalization. Selain ditujukan untuk menyamakan rentang nilai antarvariabel penelitian, proses normalisasi juga digunakan untuk mempercepat pemrosesan data selama masa *training* dan *testing* (Li et al., 2020). Contoh hasil normalisasi data untuk variabel jumlah distribusi, produksi, kebocoran, pembelian, dan pelanggan air disajikan pada Tabel 3.

Tabel 3. Contoh Hasil Normalisasi Data

No	Bulan	Produksi Air (X1)	Kebocoran Air (X2)	Pembelian Air (X3)	Pelanggan (X4)	Distribusi Air (Y)
1	Jan-15	0.69135	0.655458	0.192486	0	0.723547
2	Feb-15	0.824515	0.44343	0.03821	0.045431	0.811536
3	Mar-15	0.502044	0.360774	0.167455	0.062008	0.505789
4	Apr-15	0.811304	0.72155	0.127171	0.081347	0.834108
5	Mei-15	0.748367	0.37257	0.069323	0.10662	0.742929
6	Jun-15	0.8123	0.503395	0	0.123504	0.786972
7	Jul-15	0.661308	0.417937	0.124678	0.141922	0.668145
8	Agu-15	0.774706	0.575222	0.225128	0.156247	0.771536
9	Sep-15	0.814971	0.545855	0.180325	0.176302	0.79884
10	Okt-15	0.760879	0.46676	0.244711	0.198506	0.707158

Tabel 3 merupakan contoh hasil proses normalisasi data pada variabel jumlah distribusi, produksi, kebocoran, pembelian, dan pelanggan air yang sudah berada pada rentang 0 hingga 1. Variabel data penelitian yang sudah berada pada rentang 0 hingga 1 kini akan melalui proses seleksi fitur untuk mengetahui variabel mana yang memiliki korelasi tertinggi untuk memprediksi jumlah distribusi air.

Seleksi Fitur dengan SFS

Proses seleksi fitur dilakukan untuk mengetahui korelasi antarvariabel yang berperan sebagai prediktor (X) dengan variabel yang diprediksi (Y). Hasil dari proses seleksi fitur dengan menggunakan metode SFS disajikan pada Tabel 4.

Tabel 4. Hasil Seleksi Fitur dengan SFS

Tingkat Kesalahan	Kombinasi Fitur
0.036225	Kebocoran air

Tingkat Kesalahan	Kombinasi Fitur
0.031069	Kebocoran air, produksi air
0.035123	Kebocoran air, produksi air, pembelian air
0.037305	Pelanggan, kebocoran air, produksi air, pembelian air

Tabel 4 merupakan hasil seleksi fitur dengan metode SFS yang mampu menunjukkan tingkat kesalahan dari kombinasi seluruh variabel (X) untuk memprediksi jumlah distribusi air (Y). Tingkat kesalahan terendah yang dihasilkan melalui kombinasi variabel produksi air (X1) dan kebocoran air (X2) sebesar 0.031069 kemudian dijadikan sebagai prediktor untuk memprediksi jumlah distribusi air (Y) di masa mendatang.

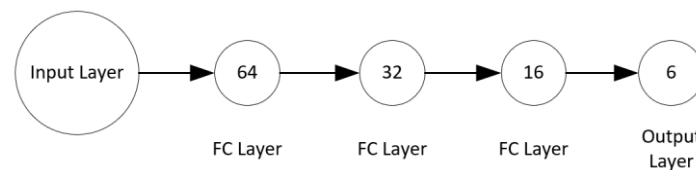
Training dan Testing Model dengan ANN serta CNN

Prediktor yang telah didapat berdasarkan metode SFS kemudian digunakan untuk membentuk model peramalan yang dibuat dengan menggunakan metode ANN serta CNN. Model peramalan jumlah distribusi air dibentuk melalui beberapa skenario *training* model yang disajikan pada Tabel 5.

Tabel 5. Skenario Pengujian Hyperparameter

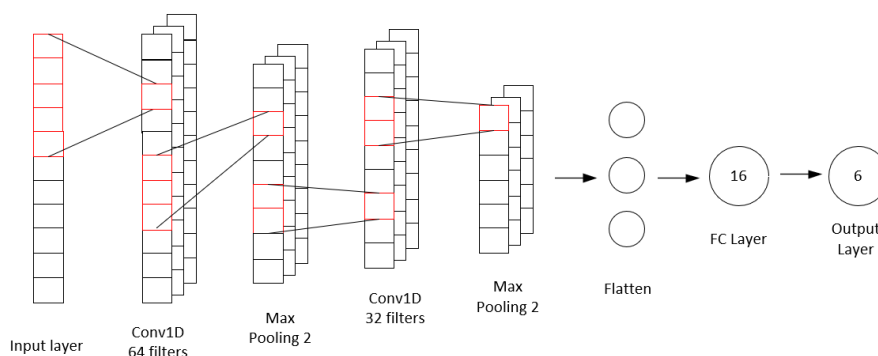
Hyperparameter	Nilai yang Diuji
Epochs	50, 100, 150
Learning Rate	0.01, 0.001, 0.0001

Tabel 5 menunjukkan *hyperparameter* berupa nilai *epochs* dan *learning rate* yang berbeda diuji dalam proses pembentukan model peramalan. Arsitektur model peramalan yang digunakan dengan metode ANN dapat dilihat pada Gambar 4.



Gambar 4. Arsitektur Model Peramalan dengan ANN

Gambar 4 merupakan kombinasi *layer* yang digunakan pada metode ANN, yaitu terdiri dari 64-32-16-6 *nodes*. *Nodes* terakhir yang berjumlah enam ditujukan untuk memprediksi jumlah distribusi air pada masa enam bulan mendatang. Arsitektur model peramalan yang digunakan dengan metode CNN dapat dilihat pada Gambar 5.



Gambar 5. Arsitektur Model Peramalan dengan CNN

Gambar 5 merupakan kombinasi *layer* yang digunakan pada metode CNN, yaitu Conv1D 64 - Max Pooling (2) - Conv1D 32 - Max Pooling (2) - *flatten layer* - 16 *nodes* - 6 *nodes*. Hasil proses *training* dan *testing* model peramalan ANN serta CNN dengan skenario pengujian yang telah ditentukan disajikan pada Tabel 6.

Tabel 6. Hasil Training dan Testing Model Peramalan Jumlah Distribusi Air

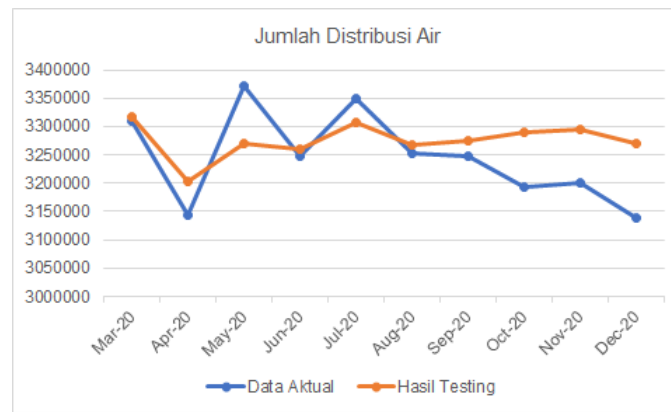
Metode	Epochs	Learning Rate	MSE	MAPE
ANN	50	0.01	0.067	2.71%
ANN	100	0.01	0.079	3.03%
ANN	150	0.01	0.109	3.23%
ANN	50	0.001	0.044	2.07%
ANN	100	0.001	0.040	2.02%
ANN	150	0.001	0.050	2.20%
ANN	50	0.0001	0.108	3.33%
ANN	100	0.0001	0.057	2.40%
ANN	150	0.0001	0.042	2.06%
CNN	50	0.01	0.087	3.26%
CNN	100	0.01	0.080	3.06%
CNN	150	0.01	0.098	3.50%
CNN	50	0.001	0.044	2.09%
CNN	100	0.001	0.042	2.07%
CNN	150	0.001	0.048	2.12%
CNN	50	0.0001	0.103	3.29%
CNN	100	0.0001	0.083	3.00%
CNN	150	0.0001	0.049	2.23%

Tabel 6 merupakan hasil proses *training* dan *testing* model peramalan dengan metode ANN dan CNN. Model peramalan paling optimal untuk memprediksi jumlah distribusi air yaitu model yang dibentuk dengan metode ANN menggunakan *epochs* 100 dan *learning rate* 0.001. Model peramalan tersebut mampu menghasilkan nilai MSE dan MAPE secara berurutan sebesar 0.040 serta 2.02%. Hasil proses *testing* model peramalan paling optimal disajikan pada Tabel 7.

Tabel 7. Perbandingan Data Aktual dan Hasil Testing Jumlah Distribusi Air

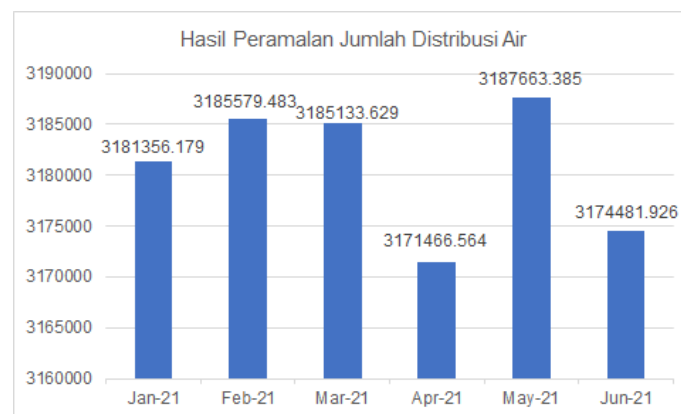
Bulan	Data Aktual (m ³)	Hasil Testing (m ³)
Mar-20	3309363	3317699
Apr-20	3143540	3202537
May-20	3371670	3269068
Jun-20	3247492	3259598
Jul-20	3349754	3308300
Aug-20	3251974	3267207
Sep-20	3247168	3274087
Oct-20	3192423	3289475
Nov-20	3199915	3294772
Dec-20	3139097	3269902

Tabel 7 menunjukkan perbandingan data aktual dan hasil *testing* jumlah distribusi air dengan menggunakan model peramalan yang paling optimal. Data tersebut kemudian divisualisasikan ke dalam bentuk grafik seperti yang terlihat pada Gambar 6.



Gambar 6. Grafik Perbandingan Data Aktual dan Hasil Testing

Gambar 6 menunjukkan bahwa pola data hasil *testing* jumlah distribusi air relatif mampu mengikuti pola data aktual jumlah distribusi air. Berdasarkan grafik yang dihasilkan, model peramalan yang dibentuk serta diuji mampu memberikan hasil peramalan yang tepat dan akurat. Model peramalan paling optimal kemudian digunakan untuk memprediksi jumlah distribusi air pada Bulan Januari 2021 hingga Juni 2021 yang hasil peramalannya dapat dilihat pada Gambar 7.



Gambar 7. Grafik Hasil Peramalan Enam Bulan Mendatang

Gambar 7 menunjukkan hasil peramalan jumlah distribusi air pada enam bulan mendatang secara berturut-turut dari Bulan Januari 2021 hingga Juni 2021 yang didapat berdasarkan model peramalan paling optimal.

Pembahasan

Air merupakan hal vital yang diperlukan dalam memenuhi kehidupan makhluk hidup setiap harinya. PDAM merupakan salah satu layanan pemerintah yang bertugas untuk menyediakan serta mendistribusikan air bersih bagi para pelanggannya. Jumlah ketersediaan air bersih bagi pelanggan tentu merupakan hal yang harus dipenuhi oleh pihak PDAM, khususnya di Kota Denpasar. Maka dari itu, salah satu cara untuk membantu mengetahui nilai jumlah distribusi air di masa mendatang adalah dengan memanfaatkan teknik peramalan menggunakan data historis yang sudah ada (Rachman, 2018; Sagheer & Kotb, 2019). Proses peramalan dilakukan dengan membandingkan metode ANN serta CNN yang juga memanfaatkan metode seleksi fitur SFS. Proses peramalan diawali dengan melakukan normalisasi data terhadap data *training* dan data *testing* yang ditujukan untuk menyamakan rentang nilai antarvariabel data penelitian dengan metode Min-Max Normalization (Altan et al., 2019; Zhu et al., 2020). Data *training* yang sudah melalui proses normalisasi kemudian digunakan untuk melakukan proses seleksi fitur guna menentukan fitur yang berpengaruh terhadap hal yang diteliti dengan menggunakan metode SFS (Haq et al., 2021). Hasil dari proses seleksi fitur kemudian digunakan pada proses *training* dan *testing* model peramalan dengan menggunakan metode ANN serta CNN. Model peramalan yang dihasilkan kemudian melalui proses evaluasi menggunakan metode MSE dan MAPE untuk meyakini bahwa hasil prediksi yang dihasilkan relevan dengan data aktual yang ada (Prayudani et al., 2019). Hasil evaluasi MSE dan MAPE yang didapat menunjukkan bahwa model peramalan jumlah distribusi yang dibentuk termasuk ke dalam kategori model peramalan yang sangat baik, sehingga tepat digunakan untuk melakukan proses

peramalan terhadap jumlah distribusi air (La Foucade et al., 2019). Berdasarkan hasil yang diperoleh, kombinasi antara metode seleksi fitur SFS dengan metode ANN menghasilkan nilai *error* yang lebih rendah jika dibandingkan dengan kombinasi metode SFS dengan CNN. Hal ini menunjukkan bahwa metode ANN lebih baik digunakan pada *dataset* penelitian karena mampu mengenali interaksi yang terjadi antara variabel yang diteliti menggunakan fitur yang diperoleh berdasarkan proses seleksi fitur dengan metode SFS (Bhakti, 2019; Trach et al., 2021). Kelebihan penelitian ini jika dibandingkan dengan penelitian (Oktafiansyah & Uperiati, 2021) terletak pada proses yang melibatkan variabel data (X) untuk memprediksi variabel data (Y) dan juga hasil yang dipaparkan lebih mudah dipahami oleh khalayak umum dengan mencantumkan persentase *error* dari model peramalan yang dihasilkan. Adapun anjuran yang dapat diberikan untuk penelitian selanjutnya, yaitu data yang digunakan perlu diperbaharui guna mendapatkan hasil peramalan yang lebih akurat serta dapat menggunakan metode seleksi fitur dan metode peramalan lainnya sebagai pembanding dengan penelitian yang telah dilakukan.

4. SIMPULAN

Model peramalan jumlah distribusi air yang dibentuk melalui model analisis yang dikembangkan menggunakan metode SFS serta ANN termasuk ke dalam kategori model peramalan yang sangat baik. Hal tersebut didukung dengan hasil pengukuran tingkat kesalahan menggunakan metode MSE dan MAPE yang bernilai relatif rendah. Melalui hasil analisis yang diperoleh, tentu mampu membuktikan bahwa model peramalan tepat digunakan untuk memprediksi jumlah distribusi air di PDAM Kota Denpasar pada masa mendatang. Hasil peramalan yang didapat secara menyeluruh tentu dapat memberikan manfaat bagi pihak PDAM Kota Denpasar dalam mempertimbangkan jumlah air yang akan disalurkan atau didistribusikan kepada para pelanggannya di masa mendatang.

5. DAFTAR PUSTAKA

- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. E., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11), e00938. <https://doi.org/10.1016/j.heliyon.2018.e00938>
- Agga, A., Abbou, A., Labbadi, M., & El Houm, Y. (2021). Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models. *Renewable Energy*, 177, 101–112. <https://doi.org/10.1016/j.renene.2021.05.095>
- Altan, A., Karasu, S., & Bekiros, S. (2019). Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. *Chaos, Solitons and Fractals*, 126, 325–336. <https://doi.org/10.1016/j.chaos.2019.07.011>
- Attoue, N., Shahrour, I., & Younes, R. (2018). Smart building: Use of the Artificial Neural Network Approach for Indoor Temperature Forecasting. *Energies*, 11(2), 1–12. <https://doi.org/10.3390/en11020395>
- Bhakti, H. D. (2019). Aplikasi Artificial Neural Network (ANN) untuk Memprediksi Masa Studi Mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik. *Eksplora Informatika*, 9(1), 88–95. <https://doi.org/10.30864/eksplora.v9i1.234>
- Castangia, M., Aliberti, A., Bottaccioli, L., Macii, E., & Patti, E. (2021). A Compound of Feature Selection Techniques to Improve Solar Radiation Forecasting. *Expert Systems with Applications*, 178. <https://doi.org/10.1016/j.eswa.2021.114979>
- Freire, P. K. de M. M., Santos, C. A. G., & Silva, G. B. L. da. (2019). Analysis of the use of discrete wavelet transforms coupled with ANN for short-term streamflow forecasting. *Applied Soft Computing Journal*, 80, 494–505. <https://doi.org/10.1016/j.asoc.2019.04.024>
- Güven, İ., & Şimşir, F. (2020). Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ANN) and support vector machines (SVM) methods. *Computers and Industrial Engineering*, 147. <https://doi.org/10.1016/j.cie.2020.106678>
- Halimi, I., Marthasari, G. I., & Azhar, Y. (2019). Prediksi Harga Emas Menggunakan Univariate Convolutional Neural Network. *Jurnal Repositor*, 1(2), 105. <https://doi.org/10.22219/repositor.v1i2.612>
- Haq, A. U., Zeb, A., Lei, Z., & Zhang, D. (2021). Forecasting Daily Stock Trend Using Multi-Filter Feature Selection and Deep Learning. *Expert Systems with Applications*, 168, 114444. <https://doi.org/10.1016/j.eswa.2020.114444>
- Harmoko, H., Triyanti, M., & Aziz, L. (2018). Eksplorasi Mikroalga di Sungai Mesat Kota Lubuklinggau. *Biodidaktika: Jurnal Biologi Dan Pembelajarannya*, 13(2), 19–23. <https://doi.org/10.30870/biodidaktika.v13i2.3366>
- Huang, C. J., Chen, Y. H., Ma, Y., & Kuo, Huan, P. (2020). Multiple-Input Deep Convolutional Neural Network

- Model for COVID-19 Forecasting in China. *MedRxiv*. <https://doi.org/10.1101/2020.03.23.20041608>
- La Foucade, A. D., Gabriel, S., Scott, E., Theodore, K., & Metivier, C. (2019). A Survey of Selected Grey Forecasting Models with Application to Medical Tourism Forecasting. *Theoretical Economics Letters*, 09(04), 1079–1092. <https://doi.org/10.4236/tel.2019.94070>
- Li, H., Zhao, W., Zhang, Y., & Zio, E. (2020). Remaining Useful Life Prediction Using Multi-Scale Deep Convolutional Neural Network. *Applied Soft Computing Journal*, 89, 106113. <https://doi.org/10.1016/j.asoc.2020.106113>
- Martín-Vázquez, R., Aler, R., & Galván, I. M. (2018). Wind Energy Forecasting at Different Time Horizons with Individual and Global Models. *IFIP Advances in Information and Communication Technology*, 519, 240–248. https://doi.org/10.1007/978-3-319-92007-8_21
- Oktafiansyah, E., & Uperiati, A. (2021). Prediksi Pendistribusian Air di Perusahaan Daerah Air Minum (PDAM) dengan Metode Adaptive Neuro Fuzzy Inference System (ANFIS) (Studi Kasus: PDAM Tirta Kepri Tanjungpinang). *Jurnal Hasil Penelitian Dan Industri Terapan*, 10(01), 32–36. <https://doi.org/10.31629/sustainable.v10i1.1404>
- Ozoegwu, C. G. (2019). Artificial Neural Network Forecast of Monthly Mean Daily Global Solar Radiation of Selected Locations Based on Time Series and Month Number. *Journal of Cleaner Production*, 216, 1–13. <https://doi.org/10.1016/j.jclepro.2019.01.096>
- Pan, H., & Zhou, H. (2020). Study on Convolutional Neural Network and Its Application in Data Mining and Sales Forecasting for E-Commerce. *Electronic Commerce Research*, 20(2), 297–320. <https://doi.org/10.1007/s10660-020-09409-0>
- Prayudani, S., Hizriadi, A., Lase, Y. Y., Fatmi, Y., & Al-Khowarizmi. (2019). Analysis Accuracy of Forecasting Measurement Technique on Random K-Nearest Neighbor (RKNN) Using MAPE and MSE. *Journal of Physics: Conference Series*, 1361(1). <https://doi.org/10.1088/1742-6596/1361/1/012089>
- Priliani, E. M., Putra, A. T., & Muslim, M. A. (2018). Forecasting Inflation Rate Using Support Vector Regression (SVR) Based Weight Attribute Particle Swarm Optimization (WAPSO). *Scientific Journal of Informatics*, 5(2), 118–127. <https://doi.org/10.15294/sji.v5i2.14613>
- Rachman, R. (2018). Penerapan Metode Moving Average Dan Exponential Smoothing Pada Peramalan Produksi Industri Garment. *Jurnal Informatika*, 5(2), 211–220. <https://doi.org/10.31311/ji.v5i2.3309>
- Sagheer, A., & Kotb, M. (2019). Time Series Forecasting of Petroleum Production using Deep LSTM Recurrent Networks. *Neurocomputing*, 323, 203–213. <https://doi.org/10.1016/j.neucom.2018.09.082>
- Sarminah, S., Pasaribu, M. B. J., & Aipassa, M. I. (2019). Pendugaan Evapotranspirasi di Lahan Agroforestri dan Lahan Terbuka Hutan Pendidikan Fakultas Kehutanan Unmul. *Agrifor: Jurnal Ilmu Pertanian Dan Kehutanan*, XVIII, 325–338. <https://doi.org/10.31293/af.v18i2.4349>
- Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., & Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Computers and Electronics in Agriculture*, 183(1432), 106036. <https://doi.org/10.1016/j.compag.2021.106036>
- Suryadibrata, A., & Chandra, D. P. (2020). Implementasi Jaringan Saraf Tiruan Backpropagation untuk Pengenalan Karakter pada Dokumen Tercetak. *Ultima Computing: Jurnal Sistem Komputer*, 11(2), 81–89. <https://doi.org/10.31937/sk.v11i2.1456>
- Syarifudin, M. A., Rini Novitasari, D. C., Marpaung, F., Wahyudi, N., Hapsari, D. P., Supriyati, E., Farida, Y., Amin, F. M., Nugraheni, R. D., Ilham, Nariswari, R., & Setiawan, F. (2021). Hotspot Prediction Using 1D Convolutional Neural Network. *Procedia Computer Science*, 179(2019), 845–853. <https://doi.org/10.1016/j.procs.2021.01.073>
- Toğa, G., Atalay, B., & Toksari, M. D. (2021). COVID-19 prevalence forecasting using Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN): Case of Turkey. *Journal of Infection and Public Health*, 14(7), 811–816. <https://doi.org/10.1016/j.jiph.2021.04.015>
- Trach, R., Trach, Y., & Lendo-siwicka, M. (2021). Using ANN to Predict the Impact of Communication Factors on the Rework Cost in Construction Projects. *Energies*, 14(14), 1–15. <https://doi.org/10.3390/en14144376>
- Xu, X., Cao, D., Zhou, Y., & Gao, J. (2020). Application of Neural Network Algorithm in Fault Diagnosis of Mechanical Intelligence. *Mechanical Systems and Signal Processing*, 141, 106625. <https://doi.org/10.1016/j.ymsp.2020.106625>
- Yan, R., Liao, J., Yang, J., Sun, W., Nong, M., & Li, F. (2021). Multi-Hour and Multi-Site Air Quality Index Forecasting in Beijing Using CNN, LSTM, CNN-LSTM, and Spatiotemporal Clustering. *Expert Systems with Applications*, 169(November 2020), 114513. <https://doi.org/10.1016/j.eswa.2020.114513>
- Zhao, X., Wei, H., Wang, H., Zhu, T., & Zhang, K. (2019). 3D-CNN-based feature extraction of ground-based

cloud images for direct normal irradiance prediction. *Solar Energy*, 181(September 2018), 510–518. <https://doi.org/10.1016/j.solener.2019.01.096>

Zhu, R., Liao, W., & Wang, Y. (2020). Short-term prediction for wind power based on temporal convolutional network. *Energy Reports*, 6, 424–429. <https://doi.org/10.1016/j.egy.2020.11.219>