Intelligent Decision Support System For Forecasting Occupancy Rate at Hotel X

Vaizal Asy’ari  
*Postgraduate of Department of Information Technology*  
*Universitas Pembangunan Nasional “Veteran” Jatim*Surabaya, Indonesia  
vaiz.asyari@gmail.com  
  
Basuki Rahmat  
*Department of Information Technology*  
*Universitas Pembangunan Nasional “Veteran” Jatim*Surabaya, Indonesia  
basukirahmat.if@upnjatim.ac.id  
  
line 1: 3rd Given Name Surname  
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCID  
  
line 1: 4th Given Name Surname  
line 2: *dept. name of organization (of Affiliation)*  
line 3: *name of organization (of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

*Abstract*— The Intelligent Decision Support System (IDSS) for Forecasting Occupancy Rates at Hotel X is designed to provide a technology-driven solution that assists hotel management in making better decisions regarding occupancy management. By utilizing machine learning models, specifically neural networks and time-series forecasting algorithms, this system analyzes historical hotel occupancy data to generate accurate predictions for future occupancy rates. The application is built using the Streamlit framework, which provides an interactive user interface, allowing users to select forecasting periods and view prediction results dynamically. The forecasting process begins with data preprocessing, including date formatting and scale adjustments, before utilizing a neural network model trained on historical occupancy data to predict future occupancy rates. Neural networks, particularly feedforward networks, are employed to capture complex patterns in the time-series data, improving prediction accuracy. The prediction results are displayed in both table and graphical formats, making it easy to visualize the comparison between historical data and forecasts. The system is also integrated into the hotel’s website using PHP, allowing hotel managers and staff to access the forecasting results directly and utilize them for more efficient decision-making regarding pricing strategies, resource allocation, and staff scheduling. With this IDSS, Hotel X can improve operational efficiency and enhance customer satisfaction through better planning and responsiveness to changes in occupancy trends. This system demonstrates significant potential in optimizing managerial processes by using accurate forecasting technology, including neural networks, which is easily accessible to users.

Keywords—hotel, occupancy rate, forecasting, decision support system, neural network

# Introduction

Business development in the Hotel Xector in Indonesia has shown a positive trend, as can be seen from the increasing occupancy rate. The occupancy rate of star classification hotels in September 2023 reached an average of 53.02%, an increase of 3.00 points compared to the occupancy rate in September 2022 of 50.02% and YTD 2023 reached an average of 49.43%, an increase of 4.04 compared to the YTD 2022 occupancy rate of 45.40% [1]. The occupancy rate of non-star classification hotels in September 2023 reached an average of 24.82%, an increase of 1.43 points compared to the occupancy rate in September 2022 of 23.39% and YTD 2023 reached an average of 23.72%, an increase of 1.52 points compared to the YTD 2022 occupancy rate of 22.21% [2]. Along with this growth, competition among hoteliers is also increasing. In-depth knowledge of the level of competition is crucial to identify opportunities, face challenges, and develop effective strategies to compete competitively.

Occupancy rate is the percentage of room occupancy rate sold [3] and is one of the indicators that can be used to see the development of hotel providers business performance in a certain period [4]. This occupancy can be calculated before the current date, which is commonly called a forecast or after the date passed [5]. By forecasting occupancy rate, the hotel can identify business opportunities or adjust room prices, identify hotel operations, and be taken into consideration for strategic decision making.

Neural networks have been around since 1943 when Warren McCulloch and Walter Pitts introduced the first neural network model calculations. This model describes the way artificial neurons can be used to process information binary. In 1950, Frank Ronseblatt continued his research by discovering a two-layer network called a perceptron. Neural network is a model with a flexible function structure, so the neural network model is rapidly developing and has been widely applied in various fields. Neural networks can be used to find solutions to problems when classical methods prove difficult or fail frequently [6].

In previous studies, several forecasting methods were used estimation of closed hotels and restaurants in Jakarta as impact of corona virus disease spread using adaptive neuro fuzzy inference system [7], forecasting the number of Demam Berdarah Dengue (DBD) patients using the fuzzy method [8], prediction of the number of visitors per period to beach attractions using triple exponential smoothing [9], classifying price range of smartphone in market using backpropagation and Learning Vector Quantification (LVQ) [10], stock price estimation using Unscented Kalman Filter (UKF) [11], forecasting of occupied rooms in the hotel using linear support vector machine [12], profitability estimation using H-Infinity and Ensemble Kalman Filter (EnKF) [13], analysis of demand and supply blood in hospital in Surabaya city using panel data regression [14], prediction of sunlight intensity using neural network and Adaptive Neuro Fuzzy Inference System (ANFIS) [15], estimation of closed hotels and restaurants as impact of Covid-19 spread using backpropagation neural network [16], neural network algorithm for breast cancer diagnosis [17], electronic nose for classifying civet coffee using Support Vector Machine (SVM), k-nearest neighbors (k-NN), and decision tree [18], forecasting agricultural products in Malang Regency using k-NN [19], forecasting average room rate using k-NN [20], forecasting occupancy rate using neural network [21]. In this study, the k-NN and neural network method with a ratio of 70%: 30% and 80%: 20% for forecasting occupancy rate at Hotel X, so that it can be used in identifying opportunities, operational implementation, and consideration in strategic decision making for management.

# Research Method

## Mind Mapping

This research adopts a structured methodology to develop an Intelligent Decision Support System (IDSS) for forecasting occupancy rates at Hotel X. The process begins with the formulation of a mind map that serves as the conceptual framework for the study. This mind map in Figure 1 below, outlines three main components: problem formulation, the proposed solution, and the expected outcomes. The problem formulation highlights the need for an automated system capable of forecasting occupancy rates, enabling more efficient operational management. The proposed solution centers on the development of a web-based system that integrates machine learning capabilities, implemented in Python, with a user-friendly interface built using PHP. The expected outcome is a system that provides accurate predictions and facilitates data-driven decision-making processes.

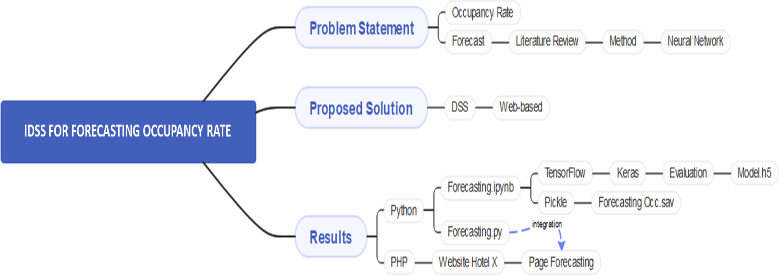


Fig. 1 Mind Mapping

## Methodology

Following the methodology (Fig. 2), the study proceeds with an extensive literature review to explore previous research, identify relevant theories, and address gaps that can be filled by the proposed system. Based on the findings, a Neural Network approach is selected as the primary forecasting method due to its proven effectiveness in handling time-series data and delivering high prediction accuracy.

Fig. 2 Methodology

The next step involves data preprocessing, where raw data is cleaned and transformed into a suitable format for model development. This process includes handling missing values, normalizing data, and splitting it into training and testing sets to prevent overfitting. Once the data is prepared, the model development phase begins by training the Neural Network algorithm using frameworks such as Keras. Model performance is then evaluated using metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to ensure reliability and accuracy.

After successful training and evaluation, the model is deployed into a web-based application. Python is used for backend computation, while PHP is employed to create a dynamic and interactive user interface. The application allows users to access real-time predictions, making it easier for stakeholders to utilize the forecasts in strategic decision-making.

Overall, this research demonstrates a systematic approach, beginning with the conceptual framework, followed by data collection, preprocessing, model development, and evaluation, culminating in the deployment of a web-based system. The final system not only delivers accurate forecasts but is also designed for scalability and ease of use, making it a reliable tool for decision support in hotel occupancy management.

## Occupancy Rate

Occupancy rate is the level of occupancy of hotel rooms calculated based on the number of rooms rented by guests and compared to the number of rooms available at a certain period [5]. The calculation of occupancy rate is shown in (1) with the result of the calculation being a percentage index measured from 0% to 100%.

(1)

The occupancy shows that out of the number of rooms available, sold. This ratio can fluctuate every day, the ratio in one month or one year is the average percentage of room occupancy sold. The highest occupancy rate is the best indicator for the hotel and as a measure of the success of hotel operations.

## Neural Network

The basics of neural networks consist of inputs, weights, processing units, and outputs. Neural networks can be applied to classifying patterns, mapping patterns obtained from inputs into new patterns in outputs, storing patterns to be recalled, mapping similar patterns, optimizing problems, and predicting. Neural networks start from preparing data for training and learning, finding neural network architecture, training and learning processes, and testing processes [22].

Neural networks can be divided into three parts called layers.

1. Input layer, responsible for receiving information, signals, features, or measurements from the external environment.
2. Hidden layers, responsible for extracting patterns related to the process or system being analyzed.
3. The output layer, responsible for producing and presenting the final tissue, results from processing by neurons in the previous layer.

A group of white circles on a black background

Description automatically generated

Fig. 3 Arrangement of neural networks in layers

Neural networks can change structures to solve problems based on internal and external information flowing through the network. Neural networks can be used to model the relationship between input and output to find patterns of data. Neurons are a basic part of the processing of a neural network. The basic shape of a neuron can be seen in Figure 4 below.

A group of white circles with black text

Description automatically generated

Fig. 4 Basic form of neurons

The weight vector (w) contains weights that connect the various parts of the network. The term "w" is used in the terminology of neural networks and is a suggestion of the expression of connections between two neurons, that is, the weight of information flowing from neuron to other neurons in the neural network. The first stage is the process of summing inputs x1, x2,... xn which is multiplied by its weight w1, w2 ,... wn expressed in equation (4):

*Net =* (*w1*. *x1* + *w2*. *x2 + w3*. *x3* + *…* + *wn*. *xn*) (4)

This concept can be written in vector notation as follows:

A black background with a black square

Description automatically generated with medium confidence

Fig. 5 Perceptron model

A threshold value of b is called a bias, which plays an important role for some neuron models and needs to be referred to as a separate neuron model parameter. Various input conditions and influences on output are required to include a nonlinear activation function f(.) in the arrangement of neurons [23]. This aims to achieve an adequate level if the input signal is small and avoid the risk of output going to inappropriate limits. Like the perceptron model in Figure 2 the output of the neuron can be expressed in terms of (5):

*y = f(net)* (5)

# Result & Discussion

The dataset used in this study is occupancy percentage data on Hotel X. The dataset used is data from April 2018 to June 2023 (63 months) can be seen in Figure 6. Furthermore, the data is split into training data and testing data. Then a test analysis was carried out with k-nearest neighbor and neural network. After that, a comparison of RMSE results from several tests of the algorithm was carried out.

A graph with blue dots and numbers

Description automatically generated

Fig. 6 Dataset occupancy rate Hotel X

## Making a Model of Neural Network

At this stage, the Hotel X dataset testing process is carried out using neural network. The Hotel X dataset that has been entered is then selected for the occupancy percentage attributes. Hotel X dataset is further divided into training data and testing data using a ratio of 80:20. At design model using neural network with one hidden layer.

One of the key parameters in training the neural network is the epoch, which defines the number of times the entire dataset is passed forward and backward through the network during the learning process. In this study, the model is trained for 150 epochs to ensure optimal learning and convergence. Each epoch allows the network to update its weights based on the error calculated during the previous iteration, gradually improving its accuracy over time. The choice of 150 epochs strikes a balance between training efficiency and performance, enabling the model to learn complex patterns within the data without overfitting or underfitting. By repeating the training process multiple times, the neural network can effectively minimize the loss function, leading to more reliable predictions. The results of this training process, including the output derived from the 150 epochs, can be observed in Figure 1, which outlines the research methodology and neural network implementation.

A screenshot of a computer

Description automatically generated

Fig. 7 Output from 150 epochs

## Neural Network Algorithm Test Analysis

The results of the test were carried out using a neural network algorithm on split data with a percentage ratio of 80:20. In testing using neural network algorithms for one hidden layer. The results of the neural network algorithm testing can be seen in Table I.

1. Test Result Using Neural Network

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hidden Layer** | **Training Data** | **Testing Data** | **RMSE** | **MAPE** |
| 1 | 80% | 20% | 0.05625455 | 0.10148834 |

The network configuration included a single hidden layer, which is designed to capture the underlying patterns and relationships within the data. The dataset was split into 80% for training and 20% for testing, ensuring that the model could learn effectively from most of the data while maintaining a portion for validation to evaluate its performance.

The performance of the model was assessed using two key evaluation metrics: Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The RMSE value obtained was 0.05625455, indicating a low level of error in the predictions, which suggests that the model performs well in capturing the trend of the data. Additionally, the MAPE was calculated to be 0.10148834 (approximately 10.15%), demonstrating that the model provides accurate predictions with minimal percentage errors relative to the actual values.

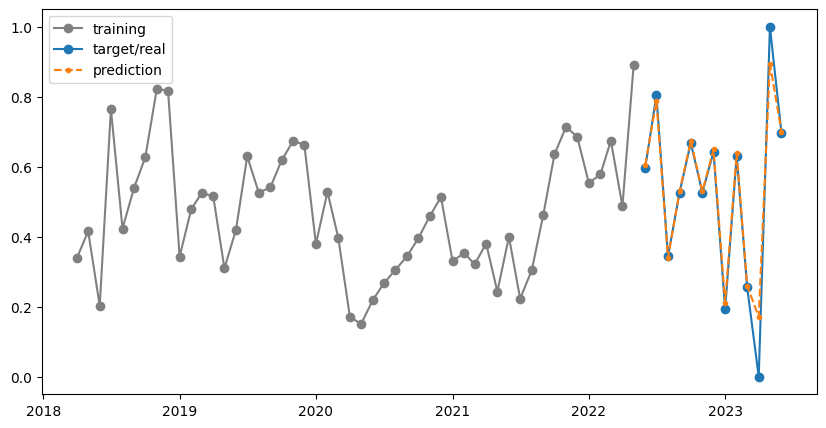


Fig. 8 Graph of testing neural network model

The graph presented in Figure 1 illustrates the performance of the Intelligent Decision Support System (IDSS) model during both training and testing phases. The training data, represented by grey dots connected with lines, shows the historical trends and patterns captured by the model during the learning process. Meanwhile, the target/real data is depicted in blue, representing the actual observed values from the testing dataset. The predictions generated by the neural network model are visualized using an orange dashed line, which closely follows the blue target points.

From the visual analysis, it is evident that the predicted values align closely with the actual values, highlighting the model's ability to track trends and fluctuations effectively. This alignment is further supported by the Root Mean Square Error (RMSE) of 0.05625455, which reflects minimal deviation between predicted and actual values, and the Mean Absolute Percentage Error (MAPE) of 10.15%, indicating a low percentage error.

The observed patterns in the graph validate the model’s robustness, especially in capturing abrupt changes and fluctuations in data, as seen in the later periods of the testing phase (2022–2023). These results demonstrate that the IDSS model can produce accurate and reliable predictions, making it a valuable tool for decision-making processes. The visual representation of results in Figure 1 reinforces the statistical evaluation metrics and provides further confidence in the system’s predictive capabilities.

## Model Iplementation

The implementation described above demonstrates the creation of a forecasting application using Streamlit, a Python framework specifically designed for building interactive and user-friendly web applications for data analysis and machine learning models. This application is focused on forecasting hotel occupancy rates, leveraging a pre-trained machine learning model stored in a serialized format and loaded using the Pickle library.

The application starts by importing necessary libraries such as Pandas for data manipulation, NumPy for numerical computations, and Matplotlib for visualizations. The dataset, stored in an Excel file (hotelx2.xlsx), is read and preprocessed to format dates and set them as indices, making it easier to handle time-series forecasting.

The user interface is built using Streamlit interactive widgets, such as the slider, which allows users to select the number of months to forecast. Based on this input, the application dynamically generates future dates for predictions. It uses the machine learning model to predict future values step-by-step by feeding the last predicted value as input for the next prediction. This iterative process generates forecasts that are then denormalized to match the original scale of occupancy rates.

The interface includes a button labeled "Forecast" that triggers the prediction process. Once predictions are computed, the application displays results in two forms: a data table and a visual plot. The data table highlights forecasted values with a distinct background color, ensuring clear differentiation between historical and predicted data. Additionally, the table formats the occupancy rates as percentages with three decimal places, enhancing readability.

A line plot visualizes the trends, where the historical data is shown in gray and the forecasted values are represented in blue. This visual representation allows users to quickly interpret patterns and assess the accuracy of predictions.

This implementation highlights Streamlit ability to create an interactive and visually appealing forecasting tool with minimal coding effort. It seamlessly integrates data preprocessing, model loading, predictions, and visualizations into a single workflow, making it suitable for users seeking a dynamic decision-support application. The approach is scalable, meaning the model can be updated, or additional features, such as downloading results, can be incorporated without major changes. The results of the application can be seen in Figure 9, where the forecasted hotel occupancy rates are plotted alongside the historical data, clearly differentiating past trends from the predicted values.

A screenshot of a computer screen

Description automatically generated

Fig. Forecasted hotel occupancy with Streamlit

The integration of the results with PHP on the hotel website is displayed in Figure 10. This integration allows for seamless forecasting output to be shown on the hotel's website, where users can interact with the forecasts, enhancing their decision-making process directly from the website interface.

A screenshot of a computer

Description automatically generated

Fig. Integration Hotel X website and application Streamlit

# Conclusion

The Intelligent Decision Support System (IDSS) for Forecasting Occupancy Rate at Hotel X offers a comprehensive and effective solution for predicting hotel occupancy trends. By integrating machine learning models, such as time-series forecasting algorithms, and leveraging data from past occupancy rates, the system provides valuable insights to aid hotel management in decision-making processes. The implementation of the model in a user-friendly web application, powered by Streamlit, allows for seamless interaction with the forecast data.

Through the use of interactive features like a dynamic forecasting slider, the system empowers users to predict occupancy rates for varying time periods. It also incorporates essential data preprocessing steps to ensure accurate predictions, while the output is displayed in a visually intuitive manner, including both numerical tables and line plots. These tools facilitate quick interpretation of trends, helping hotel management to optimize staffing, pricing strategies, and resource allocation. The integration of this forecasting system into the hotel's website, alongside the use of PHP for displaying the results, further enhances the practical application of the IDSS. Hotel managers and staff can easily access and utilize the forecast data directly from the website interface, ensuring that they are equipped with timely and relevant information to make informed decisions.

The IDSS for forecasting occupancy rates at Hotel X not only streamlines the forecasting process but also provides a powerful tool for proactive management. By enabling more accurate predictions and improving decision-making capabilities, the system contributes to increased operational efficiency, customer satisfaction, and overall profitability for Hotel X.

##### References

[1] Pusat Data dan Sistem Informasi Kemenparekraf / Baparekraf, “Tingkat Penghunian Kamar Hotel Bintang Tahun 2023,” Nov. 2023.

[2] Pusat Data dan Sistem Informasi Kemenparekraf/Baparekraf, “Tingkat Penghunian Kamar Hotel Non Bintang Tahun 2023,” kemenparekraf.go.id. Accessed: Sep. 03, 2023. [Online]. Available: https://kemenparekraf.go.id/direktori-statistik/tingkat-penghunian-kamar-hotel-non-bintang-tahun-2023

[3] B. Hermawan, *Analisis Operasional Hotel*. Pekalongan: NEM, 2022.

[4] BPS-Statistics Indonesia, “Occupancy Rate of Hotel Room 2022,” Jakarta, May 2023.

[5] Sofiani and I. S. Djunaid, *Manajemen Kantor Depan Hotel*. Yogyakarta: Bintang Semesta Media, 2023.

[6] S. Chakraverty and S. K. Jeswal, *Applied Artificial Neural Network Methods for Engineers and Scientists*. WORLD SCIENTIFIC, 2021.

[7] M. Y. Anshori *et al.*, “Estimation of closed hotels and restaurants in Jakarta as impact of corona virus disease spread using adaptive neuro fuzzy inference system,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 2, pp. 462–472, Jun. 2022.

[8] F. S. Rini, T. D. Wulan, and T. Herlambang, “Forecasting the Number of Demam Berdarah Dengue (DBD) Patients Using the Fuzzy Method at the Siwalankerto Public Health Center,” in *AIP Conference Proceedings*, American Institute of Physics Inc., Jan. 2023.

[9] A. Rahim, A. D. Rahajoe, and M. Mahaputra, “Prediksi Jumlah Pengunjung Perperiode Terhadap Tempat Wisata Pantai Menggunakan Triple Exponential Smoothing (Studi Kasus Pantai Gili Labak Sumenep),” *JIFTI-Jurnal Ilmiah Teknologi Informasi dan Robotika*, vol. 3, no. 2, pp. 39–43, Dec. 2021.

[10] M. Y. Anshori, D. Rahmalia, and T. Herlambang, “Comparison Backpropagation (BP) and Learning Vector Quantification (LVQ) on classifying price range of smartphone in market,” *J Phys Conf Ser*, vol. 1836, no. 1, p. 012040, Mar. 2021.

[11] D. F. Karya, P. Katias, T. Herlambang, and D. Rahmalia, “Development of Unscented Kalman Filter Algorithm for stock price estimation,” *J Phys Conf Ser*, vol. 1211, p. 012031, Apr. 2019.

[12] M. Y. Anshori, I. H. Santoso, T. Herlambang, D. Rahmalia, K. Oktafianto, and P. Katias, “Forecasting of Occupied Rooms in the Hotel Using Linear Support Vector Machine,” 2023.

[13] M. Y. Anshori, T. Herlambang, P. Katias, F. A. Susanto, and R. R. Rasyid, “Profitability estimation of XYZ company using H-infinity and Ensemble Kalman Filter,” in *The 5th International Conference of Combinatorics, Graph Theory, and Network Topology (ICCGANT 2021)*, IOP Publishing Ltd, Jan. 2021.

[14] A. Muhith, I. H. Susanto, D. Rahmalia, D. Adzkiya, and T. Herlambang, “The Analysis of Demand and Supply of Blood in Hospital in Surabaya City Using Panel Data Regression,” *Nonlinear Dynamics and Systems Theory*, vol. 22, no. 5, pp. 550–560, 2022.

[15] D. Rahmalia *et al.*, “Comparison between Neural Network (NN) and Adaptive Neuro Fuzzy Inference System (ANFIS) on sunlight intensity prediction based on air temperature and humidity,” *J Phys Conf Ser*, vol. 1538, no. 1, p. 012044, May 2020.

[16] F. A. Susanto *et al.*, “Estimation of Closed Hotels and Restaurants in Jakarta as Impact of Corona Virus Disease (Covid-19) Spread Using Backpropagation Neural Network,” *Nonlinear Dynamics and Systems Theory*, vol. 22, no. 4, pp. 457–467, 2022.

[17] F. S. Nugraha, M. J. Shidiq, and S. Rahayu, “Analisis Algoritma Klasifikasi Neural Network Untuk Diagnosis Penyakit Kanker Payudara,” *Jurnal Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 149–156, Aug. 2019.

[18] D. B. Magfira *et al.*, “Electronic Nose for Classifying Civet Coffee and Non-Civet Coffee,” 2023.

[19] P. Andrean, “Penerapan Metode K-NN Untuk Memprediksi Hasil Pertanian di Kabupaten Malang,” *Jurnal Mahasiswa Teknik Informatika*, vol. 3, no. 1, pp. 235–242, Mar. 2019.

[20] V. Asy’ari, M. Y. Anshori, T. Herlambang, I. W. Farid, D. Fidita Karya, and M. Adinugroho, “Forecasting average room rate using k-nearest neighbor at Hotel X,” in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)*, IEEE, Nov. 2023, pp. 496–500.

[21] M. Y. Anshori, V. Asy’Ari, T. Herlambang, and I. W. Farid, “Forecasting occupancy rate using neural network at Hotel R,” in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)*, IEEE, Nov. 2023, pp. 347–351.

[22] Imam Robandi, *ARTIFICIAL INTELLIGENCE-Mengupas Rekayasa Kecerdasan Tiruan*. Yogyakarta: Andi Offset, 2021.

[23] D. Kriesel, *A Brief Introduction to Neural Networks*. Germany: ZETA2-EN, 2005. [Online]. Available: http://www.dkriesel.com/en/science/neural\_networkshttp://www.dkriesel.com/en/tech/snipe