International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

AI Intelligence-based Gender Classification using Biometric- Digital Signature Feature Extraction methods

S. Siraj[1], Dr. Prem Kumar Singuluri[2], M. Rudra Kumar*[3]

Submitted: 06/06/2022 **Accepted:** 10/09/2022

Abstract: A person's signature can reveal much about their health, career path, and current state of mind. From a biometrics point of view, a person's gender is a demographic category like a soft trait, while a Handwritten Signature is a behavioural trait. Numerous fields, including forensics and psychology, have alluded to the possibility of gender classification based on handwritten signatures. Feminist aesthetics can be found in works written by men with a high degree of intraclass variation and vice versa. This provides evidence for using a signature to determine a person's gender. Extraction of numerical features from male and female dynamic trademark samples forms the basis of the proposed method. Five hundred thirty-five people of varying ages were surveyed. These signature examples were then transformed into numerical attributes, yielding more than 60 signature features for each dataset. Six distinct Machine Learning approaches were used in the experiments; Overall, these techniques achieved an accuracy of 78% (KNN), 83% (LR), 73% (Poly kernel-SVM), and 51% (RBF kernel in SVM). In contrast, a Poly kernel trained with cross-validation achieved 85% (SVM), 91% (DT), 97% (RF) and 98% with Deep Neural nets. In summary, deep neural networks performed best, followed closely by RF.

Keywords: *Machine Learning Algorithms, Training, Testing, Biometrics, Modeling.*

1. Introduction

Knowing an individual's signature can reveal their health, career interests, emotional state, and gender. Graphology is the name given to this study. There is a lack of international research in the field of graphology. Since these businesses see no value in learning graphology, no educational institutions offer courses in the field. Despite these obstacles, research on signatures has revealed some interesting and sometimes overlooked details. Biometric security research and related fields have spent the past decade trying to piggyback on people's unique identities. Biometrics is a person's unique physical and behavioural characteristics, such as iris, face, thumbprint, hand geometry, gait, voice, keystroke dynamics, and signature, used for identification and authentication. In recent years, monitoring technologies have been able to identify human behaviour and the social interventions individuals make when they are part of a group. Biometrics are not only useful for identification purposes but also reveal information about the individual's gender, age, national origin, and mental state. Selecting biometrics for use in various contexts requires psychological, environmental, physical, and other factors. Signatures written by hand are an advanced behavioural biometric modality that, in this specific instance, activates brain regions involved in the conception and control of actions typically associated with men. Each person's writing is distinctive because of the individual's developed consciousness of the gesture. Numerous studies have been conducted over the years about using a person's handwriting for demographic identification. There are two main types of signature collections: static (collected in a non-online setting) and dynamic (collected via the internet). In the static case, signatures are obtained offline by scanning manuscripts. Variables (X and Y coordinates, Pace, Stress, Time-series data, Angular displacement, Pen Up & down) are set in the signatures to capture the characteristics of an online handwritten signature [4]. Using the dynamical system in surveillance sciences and many other domain names were once thought impossible but has become inevitable as acquisition technology has advanced significantly. Soft characters like a person's gender, age, handedness, personality, etc. can all be analyzed by examining their digital handwritten signature, which includes natural behavioural characteristics of the individual and is stored in the cloud.

Forensic document analysis, including suicide notes, threatening letters, malevolent handwritten messages, property or rented property agreements, and other documents where identifying the person is essential necessitates gender determination. There have been many

[1]*M.Tech Student, Department of Computer Science and Engineering, G. Pullaiah College of Engineering and Technology, (Autonomous), Kurnool Email: sirajsiraj1998@gmail.com*
[2] *Sr. Professor and Dean Innovations, Department of Computer Science and Engineering, G. Pullaiah College of Engineering and Technology, Kurnool. Email: dean@gpcet.ac.in*
[3] *M. Rudra Kumar, Professor, Department of CSE, GPCET, Kurnool. Email: mrudrakumar@gmail.com*

* *Corresponding Author Email: mrudrakumar@gmail.com*

cases where the police were misled by flimsy evidence deliberately fabricated to implicate an innocent party. Forged handwritten signatures on the property and multiple checks are executed even in individual economic cases. According to the research, there are widely held beliefs about the handwriting abilities of men and women. Women's signatures are more aesthetically pleasing, consistent, and easy to read than men's. Characteristics of the individual can be determined in each case by comparing the signer's handwriting samples. As a result, the current research proposes a method of identifying a person's gender based on their online signature using various machine learning techniques [1].

2. Background and Literature

Researchers have recently demonstrated the system's efficacy by analyzing the difficulties of using internet electronic signatures for gender classification in today's building designs [8]. A person's emotional condition can be deduced from their online handwriting and their signature. There were 804 handwritten names and 8040 handwritten signatures in the database and other personal details about each person. Classifications can be made using KNN and Random Forest classifiers [12]. Both handwriting 55-58% and signatures 45-50% are recognized with high accuracy. The MRMR (Minimum Redundancy Maximum Relevancy) sequential forward selection algorithm is used for invariant feature-based age and sexual identity handwritten signature systems. In total, there are 50 unique users in the database, representing a wide range of ages and genders, as represented by every 40 signatures. A global classifier categorizes the features; the resulting error rate is 6.25 percent for sex classification and 2.6 percent for age classification [14]. Using a Deep Neural Network, the authors suggest they can identify authors of multiple languages writing online. The ADAB and IBM UB 1 handwriting standard databases are both available online and were used alongside fuzzy-based elementary perception codes and feature extraction techniques based on the Extended Beta-Elliptic model. The IBM UB 1 and ADAB datasets are accurate to within 2.5%, with the former being 96.90% and the latter 98.25%. The authors propose developing a system that can automatically determine a person's gender and dominant hand based on their handwriting and then apply that information to their online profiles. The experimentation makes use of IAMOnDB, a sizable online handwriting database. Support Vector Machine and Gaussian Mixture Model extract and categorize 29 unique features. The GMM model achieves 67.06 percent accuracy when classifying males and 84.66 percent when classifying females. Age-related changes in a person's handwriting can lead to signature variations. In the related literature, we find a small number of handwritten signature datasets that include gender annotations and datasets based on online handwriting. Currently, there is a shortage of online signature recognition system development. The gender classification of the online signature seems to have received very little attention in previous works. This research will hopefully result in the development a system that can recognize male and female handwritten signatures across a wide age range [6].

This work's contribution will be

A. The creation of a Multi-feature machine learning framework to determine a person's gender from their signature.

B. Classification methods are analyzed using a combination of Feed-Forward Deep Neural Network, k-Nearest Neighbor, Support Vector Machines, Decision Tree, Random Forest, and Logistic Regression research [23].

C. Producing a dataset in real-time that includes a variety of demographic annotations.

3. Methods

3.1. Signature-based gender categorization

Rapid advancements in machine learning and pattern recognition made widespread use of interpersonal interactions with systems commonplace [1]. A person's handwritten signature can be used as a biometric indicator of personality and health because it is unique to that person and can be quantified. Static or offline signatures are one type, and online or dynamic signatures are another. The scanned document is used to obtain static information. When using a digital signature device, a lot of useful information can be mined from the signature while it is being written online. Many data are stored, including X and Y coordinates, shape, angle, distance, orientation, stroke, velocity, pressure, angle, etc.

In most cases, its precision exceeds that of offline signature data. There is a powerful and efficient biometrics technique for determining a person's gender from an analysis of their signature data [19]. It is necessary to compare the two types of signatures to determine if a person is male or female based on their signature [4].

3.2. Machine learning-based approach to Gender Classification

In recent years, automated computer-driven image recognition systems have seen widespread use by human subjects[13]. It is possible to identify and authenticate users with the help of a variety of algorithmic systems. The branch of AI known as "Machine Learning" ML is a relatively recent development. Many images, objects, and simulated situations are used to teach these models how to complete the tasks. They can use the trained data to identify common objects, but this data is insufficient for gender classification. Data scientists often begin using a particular machine learning technique by using a properly classified dataset. Primarily, formulate an overarching principle concerning inputs and outcomes. ML models are crucial to the development of better biometric security. Embedded ML in biometrics allows for the automation of previously laborious tasks like a verification and one-to-many identification [20]. To develop and test an effective ML model, a massive dataset is required. Effective frameworks for ML-based classification and prediction algorithms have been proposed by numerous authors [17].

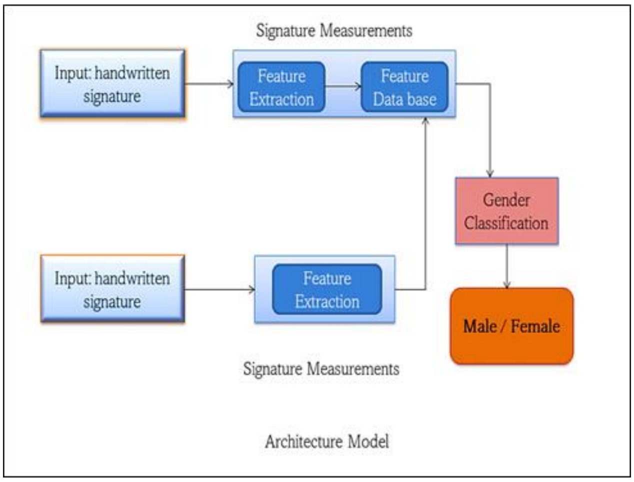

Figure 1- Proposed Architecture for the Problem statement.

3.3. Proposed Methodology

The proposed work is depicted as a block diagram in Figure 1. The proposed system model for gender categorization is shown in Figure 1.

3.3.1. Dataset Description

Figures According to the cited research, only a few minimally representative standard datasets containing only basic demographic information are currently available. Because of this, the dataset of own signatures used in the present study was collected and analyzed using the individual's preferred orientation features of their signature. Signature samples are gathered from people of varying ages and feature writing systems other than English, such as Indian languages. Anyone who was taught how to sign on a digital device and given background information on why signature samples were being collected. The acceptance and implementation of professional conduct codes and moral values are ensured by obtaining a signed consent form from each participant. Signature Device (Wacom STU-540) samples were recorded from 10 different people. They also took note of the participants' sexes and ages. 535 male and female volunteers signed the document (282 males and 253 females). Using the Secugen-Hamster Biometric device, ten fingerprint samples and a signature sample are taken from each volunteer. Ten images of each volunteer's face are taken from various angles and distances with a Nikon D3400 camera [24]. The Medplus Pharmacy's Blood Pressure Monitor, powered by a USB cable, takes the subject's systolic and diastolic readings and pulse rate from each sample [18]. Each signature sample was accompanied by an a.CSV file containing 70 system-generated features. We have zeroed in on the signature's statistical characteristics in this article. The.CSV file contains the subject's gender, age, blood pressure readings, and heart rate. As a result, there are 70 features in total. An individual's CSV file. Labels based on gender have been applied in the present study's implementation, with other domains deemed training features [4].

3.3.2. Feature Extraction

Feature extraction is a crucial step in classifying individuals based on their gender. This goal includes reducing the workforce needed to describe massive data troves.

The entire model ran on Anaconda's GUI distribution that works with various implementation environments. Jupiter notebook, Orange tool used in this context. Important features are shown in the word cloud[22].

Figure 2- Main Features

3.4. Algorithms for Classification machines and deep Learning algorithms

3.4.1. K-Nearest Neighbour k-NN - Classifier

Data Classifying data points in a multidimensional space according to a variety of distance measures for a predetermined goal is the job of the k-NN algorithm. Classification in this work will be accomplished using k-NN with an appropriate K-value that reveals k closest class label without any labels. The experiment uses an empirical cutoff of K=8 city blocks. Feature vectors are sorted into categories using k-similarity NN's measures [2]. The k-NN algorithm, a non-parametric algorithm, may find the smallest d distance between the jth training example M and the jth checking pattern N using the given [3].

3.4.2. Logistic Regression

In this sense, LR represents the broader class of classification methods that rely on predetermined classes as input. The probability of a certain outcome response variable is used to calculate predicted rates in the simplified classification model is a logistic regression classification model [7]. As the classification develops, predictions about the dependence between the labels (Y) and the features (X) are made.

3.4.3. Decision Tree

The decision tree is a well-known and frequently employed supervised learning technique for problems with two possible outcomes. The decision tree compares the data to the final estimate of the desired outcome. DT has performed a recursive split on a dataset according to the criteria, reducing it to a set of leaf nodes [5]. The training data causes it to expand from the initial root node. The Gini Index and the Entropy Index as criteria in many contexts:

The probability that a Pj is the fraction of class j nodes in state c, where E is an instance or sample that necessitates the formation of a node. The Decision Tree starts with n training samples drawn randomly from the feature dataset and then uses those samples to populate each node [21]. Continue this process until there is only one sample in each node of the same class. Using the Gini Index and the Entropy Index, it randomly selects two child nodes containing the most informative variables and features [15].

3.4.4. Support Vector Machine

SVM is a binary classifier technique that employs the boundary of a multidimensional feature vector, as well as classes from the training set and labels for those classes. Using a hyper-plane, the SVM algorithm can differentiate between training patterns that are statistically very similar to one another [9].

3.4.5. Random Forest

The Random Forest (RF) algorithm pools the results of multiple classifiers built on a tree model. Its prediction model performance is excellent and can deal with large datasets and high dimensions. During the decision tree construction process, it uses the Gini index to pick n randomly selected features and arrive at the best possible solution. New observations are nourished to the classification trees to predict the desired value of new data instances. To measure the efficacy of a classification tree, we look at how many instances of a class it correctly predicts. Class labels are returned based on which predicted class received the most votes.

3.4.6. Deep Neural Network

Despite being the most basic form of the artificial neural network, FF-DNNs find widespread use in various machine learning contexts. From the input nodes to the output units and through any intermediary nodes, all data flows forward (if any). Through trial and error, we determine the optimal combination of input, concealed, and output layers. During training, the number of input nodes will be proportional to the number of characteristics contained in the data input. Information, plus controls, weights, and a potential bias, travels down the network via a hidden layer. Finally, the activation functions help anchor the output sent to the output layer [25],[26],[27],[28]. The following is a high-level description of the proposed method's algorithm:
Firstly, a gender-based algorithm uses different machine learning methods for signature classification.

A.CSV file containing signature samples is the input.
B. Classification by gender is the result.

1. The first stage requires accessing the signature database and inputting its features and labels.
2. As a second step, use this dataset for feature engineering projects.
3. The third step is to use the following machine learning methods to categorize each sample according to gender.
Examples:
   a. Logistic regression,
   b. KNN,
   c. Support vector machine,
   d. Decision tree,
   e. Random forest,
   f. Feed-forward Artificial neural network with deep layers
4. Use criterion-based evaluations to contrast the efficacy of different approaches.
5. Stop

Process Flow Chart:
AI – ML techniques for Gender Classification using Biometrics – Signature Feature extractions

- Input: Signature samples in a .CSV file.
- Output: Gender classification

Input features and labels from the signature database
Perform feature engineering tasks on this dataset
Perform the classification of each sample into Gender classes
Machine Learning Techniques
Logistic regression,
KNN,
SVM,
Decision tree,
Random forest,

Feed-forward Deep neural network

Compare the results of each technique using evaluation metrics

Step 5: Tabulate the results

Step 6: Stop

4. Results and Discussions

4.1. Data splitting and model training

The primary goal of this investigation is to establish a gender classification system that uses distinctive physical characteristics. KNN, SVM, LR, DT, RF, and FF-DNN are some well-known ML algorithms applied to the classification problem [10]. Any machine learning (ML) algorithm is only as good as the model it was trained on. All variables are trained and tested using Train x and Test x, and the training of labels using Train y and Test y. For training and testing the model, Every piece of data is split 80/20 between training and testing. Before the model assessment, it was never exposed to the test dataset [11]. Test and training samples are listed. To prevent model over-fitting during training, k-fold cross-validation was also demonstrated in the study.

4.2. Hyperparameter Tuning

Practically all ML model architectures preclude training hyperparameters via data. However, the model's overall performance can suffer if the Hyperparameters are ill-adjusted. This is why tuning hyperparameters is so important. The accuracy of the resulting model is heavily dependent on the value of k, which is pivotal to the KNN algorithm. By default, the Sklearn library sets the value of 'k' to 5. In this investigation, we developed a bit of Python code to test the precision of the KNN model for k values between 1 and 20. After running the model on training data, the optimal value of 'k' was chosen. It follows that 'k' = 8 improved precisions significantly. Improving the accuracy of a support vector machine (SVM) algorithm by adding kernels increases the model's adaptability. The accuracy was determined by splitting the kernel into two groups. Both kernels applied the decomposition technique, and RBF and poly were selected. When the initial results weren't promising, we ran the model with scores on a cross-validation grid, determining the hyper-parameters for each subgroup of the entire dataset.

The DT classification model used the Gini and Entropy (accuracy=83%) criterion parameters. Adjusting the max depth allowed a maximum of three levels of nesting. Gini and Entropy variations produced the same level of accuracy (97%) when max-leaf nodes=100 and max features=13 were used. The RF uses the following Hyperparameters: bootstrap=True, class weight=none, criterion='gini,' max depth=none, max features='auto,' max-leaf nodes=none, min impurity decrease=0.0, min impurity split=none, min samples leaf=1, score=False, random One input layer (32 neurons, input dim=63, relu activation), two hidden layers (64 and 128 neurons, relu activation), and one output units (two neurons, activation="SoftMax") made up the sequential model used for FF-DNN. The loss element was sparse categorical cross-entropy with SoftMax activation on the final layer. Metrics were determined by calculating the "accuracy" measure, and the Adam ant colony optimization algorithm was used to handle the time complexity. The model was run for 100 epochs, with the 87th epoch yielding the highest accuracy (98%) [16].

4.3. Result Analysis

Different performance criteria, there are stand metrics used to evaluate the outcomes.

In True Positive (TP) cases, the ML model accurately forecasted the test piece into its true gender label, e.g., a female sign in the test sample is forecasted with the label 'female'. The second type, True Negative (TN), occurs when an attest sample is accurately rejected by the model based on a decisive label (A male signature is not predicted as a female signature). , A false positive (FP) occurs when an inaccurate label is anticipated for a test sample (e.g., a male sign is predicted for a female sample). Incorrectly matching a predicted sample to its old label (i.e., failing to recognize a female signature with the label "female") is known as a false negative (FN). v) Recall quantifies the fraction of class 0 samples for which the model made an accurate prediction based on the total number of such examples. vi) Accuracy measures how often a model properly identifies a sample belonging to class 0 relative to the number of samples belonging to class 0. The F-Score takes the mean of the recall and precision scores. iv)Accuracy is the proportion of

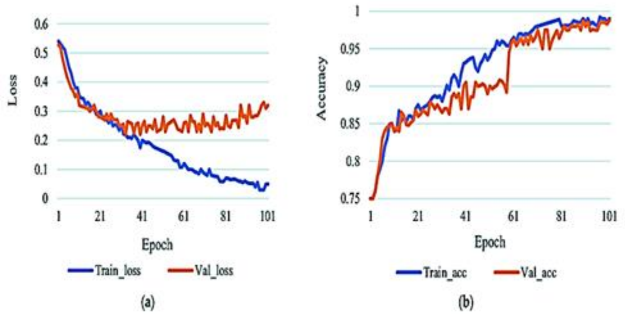

Figure 3- Accuracy Scores

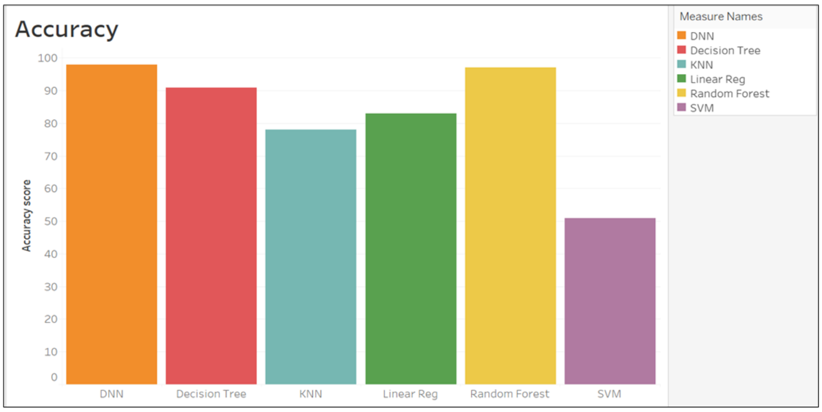

Figure 4- Accuracy Scores

Total samples that the model properly classifies. The weighted mean of each R, P, and F for every projected class is used to determine the overall average score, ignoring the relative importance of the labels. Each categorized sample's recall, accuracy, and f-score are added together and then multiplied by the total sample count for that classification class to arrive at a weighted average score. The weighted score is like the macro score but also considers the distribution of labels. x) Regardless of the model's prediction for each class, the micro average will consider the total TP, FP, and FN. All ML method results on the evaluation dataset are summarized.

The dataset is typically divided into training and test data in the traditional split approach. The reliability is based on the first set of tests. Until the task is completed in its entirety, the user has no way of knowing how precise it is. If the accuracy is poor, the user must adjust the data and run the experiment again. To eliminate this possibility, we first run the model on the training data to ensure its accuracy and then on the test data to verify its performance. Cross-validation is a technique for evaluating a model's accuracy by repeatedly running it on subsets of the training data 10% - 20% of the total. Once confidence in the validation is strong, the model is applied to the last test data set. This process is performed many times for different data folds, and the last step is determining each data fold's average accuracy. The authors have employed a cross-validation parameter to estimate the precision of the actual forecast obtained from each selected technique, adding another layer of verification to the ML models' output. When using k=20 cross-validation, the dataset is divided into ten subsets.

Furthermore, 1-20 folds are selected randomly and utilized as a training dataset, with the final fold being kept for the test set. After the split has been repeated ten times, the cross Val score statistic is used to determine the individual folds' overall accuracy. The cross Val predicts function is then used to calculate the mean score across all ten folds. Each fold's cross-validation score is listed in Each model's total accuracy is graphically represented in Figures 3, 4 and 5.

Predicted Label

True Label

Class	Male	Female
Male	1	0
Female	0.0367	0.963

Figure 5- Confusion Scores

5. Conclusion

Before identifying a certain person, there hasn't been much study on pre-processing signature photos, which is a significant limitation. Using a convolutional neural network (CNN) model, we could have encoded the extracting features task into the subjects' signatures. A.CSV file could have been generated using the method, with data consisting of sensor, match, feature, and decision level variables. Aside from that, an autoencoder could have been used to omit the features by referencing the latent space representation. Using fewer features here (rather than 66 features) would have been preferable. To complete the data augmentation process, no extra methods are employed. The gathered features were anticipated to prove useful in the classification process. Assisting with segment By augmenting the data, we may have trained new algorithms based on ML models that showed fascinating patterns, allowing us to classify other subcategories, such as identifying neurological illnesses.

The following are examples of what needs to be done next for this research project:

A. Use machine learning (ML) algorithms to analyze these distinctive characteristics of any human neurological disorder.

B. Examine the emerging multi-class signature-based challenges, such as semantic language identification and age prediction with signature features.

C. Increase the dataset size by gathering additional samples of handwritten signatures to boost the obtained results with specific demographic data.

In this study, the authors use the numerical functionality of the signature dataset to perform binary classification for gender identification, using a total of six different ML algorithms. The proposed comparison uses fast training, high accuracy, and comparatively simple models. The DNN had the highest overall classification accuracy (98%) among the selected six algorithms.

References

[1]. D. M. Abdullah and A. M. Abdulazeez, "Machine Learning Applications based on SVM Classification A Review," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 81–90, 2021.

[2]. T. N. Lakshmi, S. Jyothi, and M. R. Kumar, "Image Encryption Algorithms Using Machine Learning and Deep Learning Techniques—A Survey," in Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough, Springer, 2021, pp. 507–515.

[3]. S. Begum, D. Chakraborty, and R. Sarkar, "Data classification using feature selection and kNN machine learning approach," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN), 2015*, pp. 811–814.

[4]. Meneses, B., E. L. Huamani, M. Yauri-Machaca, J. Meneses-Claudio, and R. Perez-Siguas. “Authentication and Anti-Duplication Security System for Visa and MasterCard Cards”. International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 7, July 2022, pp. 01-05, doi:10.17762/ijritcc.v10i7.5558.

[5]. Y. Serdouk, H. Nemmour, and Y. Chibani, "New offline handwritten signature verification method based on artificial immune recognition system," *Expert Syst. Appl.*, vol. 51, pp. 186–194, 2016.

[6]. M. Somvanshi, P. Chavan, S. Tambade, and S. V Shinde, "A review of machine learning techniques using decision tree and support vector machine," in *2016 international conference on computing communication control and automation (ICCUBEA), 2016*, pp. 1–7.

[7]. X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feed-forward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010*, pp. 249–256.

[8]. A. I. Schein and L. H. Ungar, "Active learning for logistic regression: an evaluation," Mach. Learn., vol. 68, no. 3, pp. 235–265, 2007.

[9]. Ghazaly, N. M. . (2022). Data Catalogue Approaches, Implementation and Adoption: A Study of Purpose of Data Catalogue. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(1), 01–04. https://doi.org/10.17762/ijfrcsce.v8i1.2063

[10]. A. Ghildiyal, S. Sharma, I. Verma, and U. Marhatta, "Age and gender predictions using artificial intelligence algorithm," in *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020*, pp. 371–375.

[11]. J.-H. Yoo, D. Hwang, and M. S. Nixon, "Gender classification in human gait using support vector machine," in *International Conference on Advanced Concepts for Intelligent Vision Systems, 2005*, pp. 138–145.

[12]. M. S. Lee, L. N. Guo, and V. E. Nambudiri, "Towards gender equity in artificial intelligence and machine learning applications in dermatology," *J. Am. Med. Informatics* Assoc., vol. 29, no. 2, pp. 400–403, 2022.

[13]. M. Rudra Kumar and V. K. Gunjan, "Machine Learning Based Solutions for Human Resource Systems Management," in *ICCCE 2021, Springer, 2022*, pp. 1239–1249.

[14]. Hermina, J. ., Karpagam, N. S. ., Deepika, P. ., Jeslet, D. S. ., & Komarasamy, D. (2022). A Novel Approach to Detect Social Distancing Among People in College Campus. International Journal of Intelligent Systems and Applications in Engineering, 10(2), 153–158. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/1823

[15]. M. R. Segal, "Machine learning benchmarks and random forest regression," 2004.

[16]. S. Hu, "Research on data acquisition algorithms based on image processing and artificial intelligence," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 34, no. 06, p. 2054016, 2020.

[17]. S. Kumar, S. Singh, and J. Kumar, "Gender classification using machine learning with the multi-feature method," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 2019*, pp. 648–653.

[18]. M. R. Kumar, V. D. Chakravarthy, T. N. Ranganatham, and K. Ramana, "Personal finance transaction index scoring using machine learning model," *Mater. Today, Proc.*, 2021.

[19]. S. Kumar, S. S. Gornale, R. Siddalingappa, and A. Mane, "Gender Classification Based on Online Signature Features using Machine Learning Techniques," *Int. J. Intell. Syst. Appl.*

Eng., vol. 10, no. 2, pp. 260–268, 2022.

[20]. D. J. Reddy and M. R. Kumar, "Crop yield prediction using machine learning algorithm," in *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021*, pp. 1466–1470.

[21]. N. A. Libre. (2021). A Discussion Platform for Enhancing Students Interaction in the Online Education. Journal of Online Engineering Education, 12(2), 07–12. Retrieved from http://onlineengineeringeducation.com/index.php/joee/article/view/49

[22]. A. Sumesh, K. Rameshkumar, K. Mohandas, and R. S. Babu, "Use of machine learning algorithms for weld quality monitoring using an acoustic signature," *Procedia Comput. Sci., vol. 50,* pp. 316–322, 2015.

[23]. K. Bobkowska, A. Janowski, M. Przyborski, and J. Szulwic, "Analysis of high resolution clouds of points as a source of biometric data," in *2016 Baltic Geodetic Congress (BGC Geomatics), 2016*, pp. 15–21.

[24]. G. Seeja, O. Reddy, K. V. R. Kumar, S. Mounika, and Others, "Internet of Things and Robotic Applications in the Industrial Automation Process," in *Innov. Ind. Internet Things Smart Fact., IGI Global, 2021*, pp. 50–64.

[25]. Degambur, L.-N., Mungur, A., Armoogum, S., & Pudaruth, S. (2022). Resource Allocation in 4G and 5G Networks: A Review. International Journal of Communication Networks and Information Security (IJCNIS), 13(3). https://doi.org/10.17762/ijcnis.v13i3.5116

[26]. S. D. Bhogaraju, K. V. R. Kumar, P. Anjaiah, J. H. Shaik, and Others, "Advanced Predictive Analytics for Control of Industrial Automation Process," in *Innov. Ind. Internet Things Smart Fact., IGI Global, 2021*, pp. 33–49.

[27]. K. Hemachandran et al., "Artificial Intelligence: A Universal Virtual Tool to Augment Tutoring in Higher Education," *Comput. Intell. Neurosci., vol. 2022*, 2022.

[28]. M. K. Satheesh and K. V. R. Kumar, "Addressing the Utilization of Popular Regression Models in Business Applications," in *Machine Learning for Business Analytics, Productivity Press*, pp. 29–43.

[29]. Meshram, A. A. ., & Abhimanyu Dutonde, A. G. . (2022). A Review of Skin Melanoma Detection Based on Machine Learning. International Journal of New Practices in Management and Engineering, 11(01), 15–23. https://doi.org/10.17762/ijnpme.v11i01.145

[30]. E. M. K. Reddy, A. Gurrala, V. B. Hasitha, and K. V. R. Kumar, "Introduction to Naive Bayes and a Review on Its Subtypes with Applications," *Bayesian Reason. Gaussian Process. Mach. Learn. Appl., pp*. 1–14, 2022.

[31]. Madapuri, Rudra Kumar, and P. C. Mahesh. "HBS-CRA: scaling the impact of change request towards fault proneness: defining a heuristic and biases scale (HBS) of change request artifacts (CRA)." Cluster Computing 22.5 (2019): 11591-11599

[32]. Rudra Kumar, M., Rashmi Pathak, and Vinit Kumar Gunjan. "Machine Learning-Based Project Resource Allocation Fitment Analysis System (ML-PRAFS)." *Computational Intelligence in Machine Learning*. Springer, Singapore, 2022. 1-14.

[33]. Kumar, M. Rudra, Rashmi Pathak, and Vinit Kumar Gunjan. "Diagnosis and Medicine Prediction for COVID-19 Using Machine Learning Approach." *Computational Intelligence in Machine Learning: Select Proceedings of ICCIML 2021* 834 (2022): 123.

[34]. Suneel, Chenna Venkata, K. Prasanna, and M. Rudra Kumar. "Frequent data partitioning using parallel mining item sets and MapReduce." International Journal of Scientific Research in Computer Science, Engineering and Information Technology 2.4 (2017)