Skin Tone Benchmark Dataset for Diabetic Foot Ulcers and Machine Learning to Discover the Salient Features

Conference Paper · July 2022

CITATIONS
3

READS
140

5 authors, including:

Masrur Sobhan
Florida International University
27 PUBLICATIONS **131** CITATIONS
SEE PROFILE

Md Abdullah Al Mamun
Florida International University
31 PUBLICATIONS **166** CITATIONS
SEE PROFILE

Anuradha Godavarty
Florida International University
149 PUBLICATIONS **1,837** CITATIONS
SEE PROFILE

Skin Tone Benchmark Dataset for Diabetic Foot Ulcers and Machine Learning to Discover the Salient Features

Masrur Sobhan[1#], Kacie Kalie[2#], Abdullah Al Mamun[1], Anuradha Godavarty[2], Ananda Mohan Mondal[1*]

[1] Machine Learning and Data Analytics Group
School of Computing and Information Sciences,
Florida International University, Miami, USA
{msobh002, mmamu009, amondal}@fiu.edu

[2] Optical Imaging Laboratory
Biomedical Engineering,
Florida International University, Miami, USA
{kkail001, godavart}@fiu.edu

*Corresponding author; #Equal contribution

Abstract. One in three people with diabetes develops diabetic foot ulcers (DFUs) during their lifetime, which is a major risk factor for amputation and mortality. Optical imaging techniques have been developed to determine the extent of oxygenation to these DFUs, which is vital to assessing DFU wounds. Herein, a machine learning approach was implemented to label different skin tones to eventually correct the tissue oxygenation maps. We developed a skin tone benchmark dataset of 9,000 samples and a machine learning framework to represent this dataset in a reduced dimension of 20 features, which could subsequently be incorporated into our smartphone-based optical imaging device developed for DFU assessments. This allows our technology to be applicable across different racial/ethnic groups of varying skin tones.

Keywords: Deep learning, Diabetes foot ulcer, Feature selection, Fitzpatrick skin tone, smartphone-based application.

1 Introduction

Diabetes is one of the major global health problems affecting public health and socio-economic progress [1]. The severity of diabetes has grown much in the last twenty years [2]. In recent days, diabetes has been considered a silent pandemic, even deadlier than COVID-19 [3]. Diabetes is responsible for a high mortality rate due to contagions, cardiovascular disease, stroke, chronic diseases, and cancer [4, 5]. Particularly diabetes foot ulcer (DFU) is one of the major and common difficulties of patients with diabetes, and it is associated with disability and morbidity along with impairment of life [6, 7]. One in three people with diabetes develops DFU sometime during their lifetime [8]. DFUs may lead to complications in diabetic people, even death [9]. Though not all physicians can treat DFUs, it is extremely necessary to have some knowledge to perform an initial evaluation by addressing three major concerns: debridement, offloading, and infection control [10]. But due to the global pandemic, COVID-19, DFU care has been disrupted [11]. So, it is necessary to think about an alternative treatment such as at-home point-of-care (POC) rather than visiting clinics. Assistive technology with artificial intelligence (AI) may play a vital role in mitigating the risks due to DFUs.

Recent studies [12, 13] showed that a smartphone-based device could measure tissue oxygenation. But it does not account for variations in skin tone or pigmentation of various racial groups with DFUs, wound tissue characteristics, and foot curvature to obtain

accurate tissue oxygenation measurement. Researchers found disparities in the incidence of DFUs among Caucasians and other races, such as African Americans, Hispanics, and Native Americans, depending on their skin tones [14]. Thus, our smartphone-based device needs to determine these skin tones and account for the melanin concentration differences during tissue oxygenation measurements. Hence, machine learning algorithms are implemented to classify the skin tones in the current study before correcting for the tissue oxygenation maps.

Here, we developed a benchmark dataset for skin tone types and a machine learning framework to identify the salient features of different skin tone types. In our future work, the developed ML framework will be incorporated into a smartphone-based device so that it can generate patient sample data annotated with both DFU and skin tone type.

The people's skin tone can be divided into six groups (I- VI), where group I is the lightest skin and group VI is the darkest, known as Fitzpatrick skin tone (FST) [15]. Classifying the skin tone using a smartphone by applying machine learning or deep learning algorithms [16] from images captured may minimize the human error of the clinicians and make the overall process automated. But an image has many pixel values, which are considered features. Processing a large number of features to classify skin tone is computationally expensive. Rather, identifying the key features, a subset of the most informative features, from the images and using only those features for classifying skin tone is more efficient than using all the features or pixels. In this study, we applied five state-of-the-art machine learning algorithms, Least Absolute Shrinkage and Selection Operation (LASSO) [17], Multi-Cluster Feature Selection (MCFS)[18], Random Forest (RF)[19], Support Vector Machine with Recursive Feature Elimination (SVM-RFE) [20], Unsupervised Discriminative Feature Selection (UDFS) [21], and a deep learning algorithm Multi-run Concrete Autoencoder (mrCAE) [22], which are capable of identifying the important set of key features. To check the importance of the identified key features, we used a machine learning algorithm, SVM, to classify skin tone types. The classification performance was used to measure the importance of the key features. Our long-term goal is to establish a smartphone-based application to report both DFU and skin types by capturing the images of DFU wounds.

2 Methodology

2.1 Dataset Generation and Processing

Fig. 1 outlined the process of dataset generation and processing for Fitzpatrick skin tone types (I-VI).

Data generation. The color images of 10 frames of Fitzpatrick scale sticker consisting of six skin tone types I-VI were taken at three ambient light levels – LOW, MEDIUM, and HIGH. Fifty regions of interest (ROIs) were randomly generated as 10×10-pixel regions for each skin type (I-VI), across each frame (1-10), and for each light level.

Data processing: Each ROI is a sample labeled with the corresponding skin type. The samples were uniformly distributed across six skin types each having 1,500 (50 ROIs/frame x 10 frames/light condition x 3 light conditions)) samples, thus generating a total of 9,000 samples. Each sample with 10×10 pixels of 3 channels (R, G, B) was flattened to make 300 features. So, the final dataset used for the feature selection algorithm consists of a matrix of 9000 samples × 300 features.

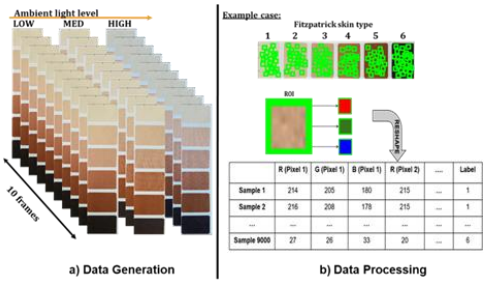

Fig. 1. Dataset generation and processing. (a) Data generation: Color images of 10 frames of FST sticker were taken at three ambient light level – LOW, MEDIUM, and HIGH. (b) Data Processing: Each ROI of 10x10 pixels represents a sample with 300 (10x10x3) features labeled with the corresponding skin tone type (I-VI).

2.2 Feature selection using machine learning and deep learning algorithms

To represent the skin tone dataset in a reduced dimension, we explored six feature selection algorithms. Fig. 2 shows the feature selection framework and evaluation of the performance of feature selection algorithms. Five algorithms were machine learning-based, including LASSO, MCFS, RF, SVM-RFE, and UDFS, and one deep learning-based, Concrete Autoencoder (CAE) [23]. LASSO, RF, and SVM-RFE are frequently used embedded feature selection algorithms. The other three – CAE, MCFS, and UDFS are unsupervised feature selection algorithms. The machine learning algorithms were implemented using the Scikit-Learn [24] framework, and CAE was implemented using a deep learning framework named Keras [25].

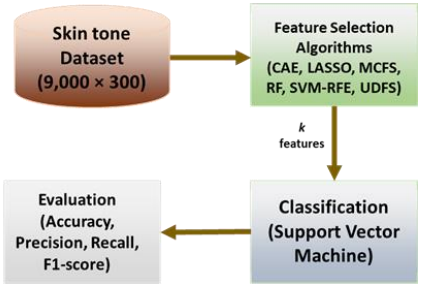

Fig. 2. Feature Selection Framework for Discovering Salient Features for Skin Tones. Input: Skin tone dataset of 9,000 samples (ROIs from images) with 300 features. Six feature selection algorithms were used for analysis. The selected features were used in classifying skin tone types to check the performance of feature selection algorithms.

Concrete Autoencoder (CAE) is a variation of Autoencoder (AE). But the main difference between CAE and AE is that CAE finds a reduced set of most informative actual features, whereas AE derives a new set of reduced features. A CAE consists of a single encoder layer, also called the feature selection layer, and decoder layers, as shown in Fig. 3.

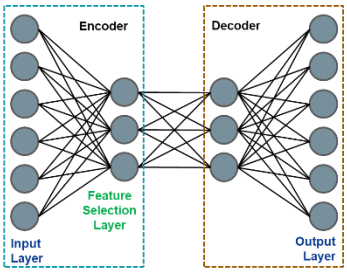

Fig. 3. Architecture of concrete autoencoder. The encoder layer is the feature selection layer. The working mechanism of concrete autoencoder is similar to autoencoder. But the difference is that feature selection layer captures the salient features whereas autoencoder captures the latent features.

The task of the feature selection layer is to find a subset of (k) actual features in a stochastic manner from the original set of features (FST skin tone dataset with 300 features in this study). On the other hand, decoder layers try to reconstruct the original features using only k number of features chosen by the feature selection layer. The detailed algorithm of this method is discussed in [23]. Temperature and mean-max probability (mean of maximum probabilities of the selected features) play a vital role in selecting the most meaningful k features from the original set. Every node in the feature selection layer assigns a probability value for each feature and selects the feature with the maximum probability. Temperature is a hyperparameter applied to the logits, which affects the final probabilities from the SoftMax output [26]. The lower the temperature, the higher the model confidence (i.e., selecting a stable subset of features), as shown in Fig. 4. On the other hand, mean-max probability helps make the model confident when the mean-max probability value is high, Fig. 4. So, the main task of these two hyperparameters is to make the model confident. One disadvantage of CAE is that it selects different meaningful k features in different runs.

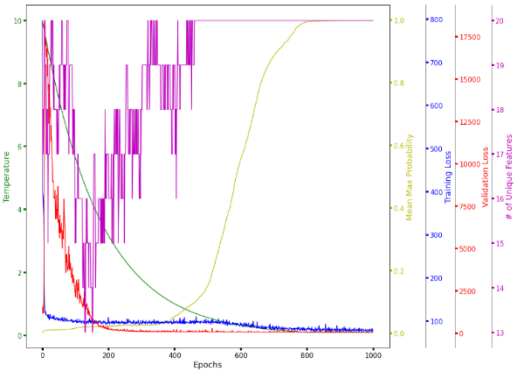

Fig. 4. Characteristic curve of concrete autoencoder. X-axis represents the number of epochs. Y-axis has various representations. Green: temperature [range: 10 to .01]. Yellow: mean-max probability [threshold >= 0.998]. Blue: training loss. Red: validation loss. Purple: number of unique features. The number of features (k) was set to 20.

Recent studies [22, 27] proposed a multi-run concrete autoencoder (mrCAE) to identify a stable set of features by running CAE multiple times and then pick up the most frequent k features, which are considered significant features. The procedure and mechanism of mrCAE are described in [22]. In this study, we used the same technique (mrCAE) to identify significant features or pixels from the original data with 300 pixels.

Hyperparameter Tuning and Characteristic Curve for CAE: Grid-search approach was used to tune the number of epochs and learning rate. We used the same value of temperatures (starting from 10 and ending at 0.1) as used in [23]. Our analysis

showed that the model with a learning rate of 0.001 produces a mean-max probability of 0.998 for the present dataset while the number of epochs is around 1,000, as shown in the characteristic curve, Fig. 4. The characteristic curve, as demonstrated by Mamun et al. [22], is a way of showing that the values of tuned hyperparameters are optimal. The characteristic curve shows that the unique number of features was initially unstable. But, with the decrease in temperature, the mean-max probability increases, meaning the model becomes more confident in selecting unique features, thus producing a stable convergence in training and validation losses.

The dataset was split into the train and validation sets according to the 80/20 ratio maintaining the distribution of classes to avoid overfitting for all the six algorithms used in this study. The training set was used to estimate the learning parameters, and the validation set was used for performance evaluation.

2.3 Feature selection

Machine learning-based feature selection algorithms (LASSO, MCFS, RF, SVM-RFE, UDFS) were run for a single time, as they selected the same set of features in each run. On the other hand, we ran the deep learning-based feature selection algorithm, mrCAE, 10 times to get a meaningful set of features since it produces different sets of features in different runs due to its stochastic nature. We considered the features significant and meaningful that appeared the most (higher frequency) in different runs [22].

2.4 Performance Evaluation of Feature Selection Algorithms

To check the importance of the features selected by six algorithms, i.e., the feature selection performance, we used the selected features to classify six skin tones using a support vector machine (SVM), Fig. 2. The hyperparameter tuning for SVM was conducted using Grid-Search. Various combinations of parameters, C (adds a penalty for each misclassified datapoint; if ‘C’ is small, decision boundary of large margin is chosen with more misclassified datapoint) and the kernels (linear, poly, RBF, or sigmoid) were used while classifying using SVM. 5-fold cross validation was used to measure the classification performance. We made sure that in each of the folds, the ratio of training and validation for each of the labels remained the same (80/20 split for each skin tone). The classification performance was evaluated using four metrics: accuracy, precision, recall, and f1-score.

3 Result

3.1 Optimal Set of Features

We ran six feature selection algorithms, including LASSO, MCFS, RF, SVM-RFE, UDFS, and mrCAE, to select five subsets of k features (k = 5, 10, 20, 30, and 40 features) from the original set of 300 features. We classified the skin tone dataset using the selected features with the support vector machine (SVM) classifier. Fig. 5 shows

the plot of classification accuracy using five subsets of features selected by each of the six feature selection algorithms. We noticed that the accuracies of all the algorithms increased with the increase of k up to 20. But the increase was insignificant with k > 20. Thus, we chose 20 as the optimal number of features.

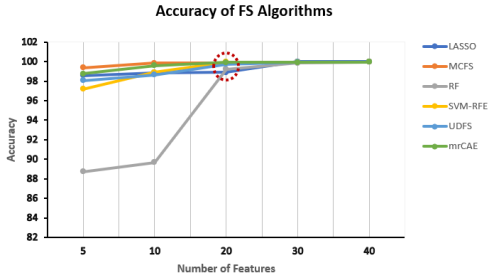

Fig. 5. Selecting the optimal set of features. Accuracy with different subsets of features selected by six feature selection algorithms – LASSO, MCFS, RF, SVM-RFE, UDFS, and mrCAE, are plotted against the number of selected features.

3.2 Evaluation Results

We employed two approaches to evaluating the significance of the optimal set of features, i.e., the performance of the feature selection algorithms. *Approach-1:* Classification performance of the selected features in classifying skin tone dataset into six classes (I-VI). *Approach-2:* Capability of the selected features in clustering skin tone dataset into different skin tone types using t-SNE [28]. The significant 20 features selected from the six algorithms were used for evaluation.

Evaluation based on classification performance. Fig. 6 shows the accuracy, precision, recall, and f1 score for LASSO, MCFS, RF, SVM-RFE, UDFS, and mrCAE. All the scores are very similar for the six algorithms. But SVM-RFE had a slightly better score than other feature selection algorithms. The maximum classification accuracy was 99.98% (SVM-RFE), and the minimum was 98.91% (LASSO). The scores suggest that the differences in the scores are not significant.

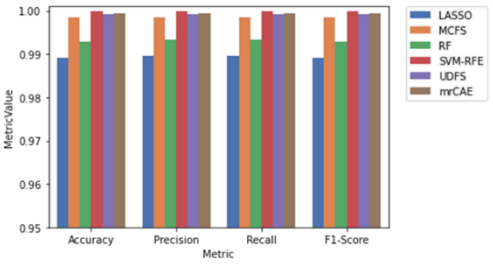

Fig. 6. Evaluation of the 20 selected features. Six set of 20 features identified from six different algorithms were used for classification. Accuracy, precision, recall and f1-score were used as evaluation metrics.

Evaluation based on t-SNE plots. We used t-SNE to cluster the samples using all the features and the selected 20 features from mrCAE. Fig. 7(a) shows distinct clusters for all the different Fitzpatrick skin tones using all the features. All skin tones have more than one cluster because of the light levels (low, medium, and high). Similar clusters were observed using the significant 20 features derived from mrCAE, Fig 7(b). We also noticed that all the six sets of 20 features derived from six algorithms produced similar clusters as the original. It is clear from the t-SNE plots that all the

algorithms could identify the most meaningful 20 features from the original data with 300 features.

It is noticeable from Fig. 7(b) that skin tones V and VI have three, and the other types have two distinct clusters. These results are due to the three different light conditions - LOW, MEDIUM, and HIGH. As skin tones, V and VI are the darkest compared to the other four (I-IV), mrCAE captured salient features of samples in all three light conditions. In the latter case, the LOW- and MEDIUM-light samples were combined into one cluster. So, it is clear that mrCAE can also identify features at a more granular level. The same behavior is observed in the case of the other five feature selection algorithms (not shown).

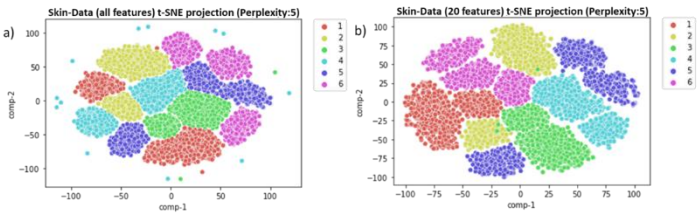

Fig. 7. t-SNE plots for clustering. (a) t-SNE plot using all the features. (b) t-SNE plot using only 20 significant features selected by mrCAE. Both sets of features (all 300 and 20 significant features) produced the similar clusters.

4 Discussion

Our hypothesis was that a small subset of salient features would be able to represent the FST skin tone dataset and classify the same into six skin tone types with high accuracy. To test the hypothesis: first, we developed a benchmark dataset for FST skin tone types; second, we developed a machine learning framework to represent the skin tone dataset using a fewer number of salient features.

We used three embedded and two unsupervised machine learning-based and one deep learning-based feature selection algorithms to identify the salient features from FST skin tone images. We ran all the algorithms once as the selected features were reproducible except CAE. As CAE produces different sets of important features in different runs, we ran CAE 10 times (called mrCAE) and identified the most frequent features as the significant ones. Ten runs of CAE with k equal to 20 (number of selected features) produced 200 features. We picked the most frequent 20 features whose frequency ranges between 6 (maximum) to 2 (minimum). Then a classifier, SVM, was applied to classify the FST skin tones using the six sets of 20 salient features. The selected 20 features produced high accuracies ranging between 98.91% (with LASSO-derived features) to 99.98% (with SVM-RFE-derived features). The reason for such high accuracies could be because the generated data was too ideal since we used stickers for FST to generate the data.

The t-SNE plots using selected 20 features showed that the two darkest skin tones (Types V and VI) produced three clusters corresponding to samples with low, medium,

and high light levels. Whereas other skin types (lighter types: I – IV) produced two clusters, one for the samples with high light level and the other for the combined samples with low and medium lights. Thus, the selected 20 features can distinguish the skin tones not only into six skin tone types but also into granular levels. These results also suggest that we should use the same light condition for generating all the data, preferably "HIGH" light condition, to ensure that six skin tones produce six distinct clusters, each representing a single skin tone.

5 Conclusion and Future Remarks

This study represents a preliminary experiment and observation of a long-term goal to establish a smartphone-based application to report DFU and skin types by capturing the images of DFU wounds. This study may help clinicians design appropriate treatment strategies for an individual patient with confidence. Using the whole images captured by a smartphone for classifying the skin tones with the same smartphone may take lots of time as the original image contains a large number of pixels or features. This issue can be tackled by using salient features (instead of all features) in classifying the skin tones. With this aim, we explored six algorithms on the 9000 FST skin tone samples with 300 features. Among the 300 features, we identified six sets of the best 20 features, which produced a near perfect (~100% accuracy) classification of six skin tone types.

In the future, we will include real patients' images and employ a larger database for in-vivo skin tone classification studies. Thus, using the indexes of these salient features could help in our future work of developing an ML-based mobile app to automate skin tone classification and related melanin corrections when analyzing tissue oxygenation maps using near-infrared optical measurements.

Acknowledgments

This work was partially supported by the NSF CAREER Award IIS 1901628 to AMM; Pilot award, NIH/NIMHD U54 MD012393 to AMM, and Coulter Med Tech Grant (BME-FIU) to AG. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

1. Lin, X., Xu, Y., Pan, X., Xu, J., Ding, Y., Sun, X., Song, X., Ren, Y., Shan, P.F.: Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Sci. Rep. 10, 1–11 (2020). https://doi.org/10.1038/s41598-020-71908-9
2. Ramachandran, A., Snehalatha, C., Shetty, A.S., Nanditha, A.: Trends in prevalence of diabetes in Asian countries. World J. Diabetes. 3, 110 (2012). https://doi.org/10.4239/WJD.V3.I6.110

3. Diabetes “Silent pandemic” 3 times more deadlier than COVID-19 | Africa Health Organisation, https://aho.org/news/diabetes-silent-pandemic-3-times-more-deadlier-than-covid-19/
4. Bragg, F., Holmes, M. V., Iona, A., Guo, Y., Du, H., Chen, Y., Bian, Z., Yang, L., Herrington, W., Bennett, D., Turnbull, I., Liu, Y., Feng, S., Chen, J., Clarke, R., Collins, R., Peto, R., Li, L., Chen, Z., Group, for the C.K.B.C.: Association Between Diabetes and Cause-Specific Mortality in Rural and Urban Areas of China. JAMA. 317, 280–289 (2017). https://doi.org/10.1001/JAMA.2016.19720
5. Policardo, L., Seghieri, G., Anichini, R., De Bellis, A., Franconi, F., Francesconi, P., Del Prato, S., Mannucci, E.: Effect of diabetes on hospitalization for ischemic stroke and related in-hospital mortality: A study in Tuscany, Italy, over years 2004-2011. Diabetes. Metab. Res. Rev. 31, 280–286 (2015). https://doi.org/10.1002/DMRR.2607
6. Oliver, T.I., Mutluoglu, M.: Diabetic Foot Ulcer. StatPearls. (2021)
7. Frykberg, R.G., Banks, J.: Management of Diabetic Foot Ulcers: A Review. Fed. Pract. 33, 16 (2016)
8. Boulton, A.J.M.: Diabetic foot disease during the COVID-19 pandemic. Med. 57, 1–9 (2021). https://doi.org/10.3390/medicina57020097
9. Yazdanpanah, L., Nasiri, M., Adarvishi, S.: Literature review on the management of diabetic foot ulcer. World J. Diabetes. 6, 37 (2015). https://doi.org/10.4239/WJD.V6.I1.37
10. Kruse, I., Edelman, S.: Evaluation and Treatment of Diabetic Foot Ulcers. Clin. Diabetes. 24, 91–93 (2006). https://doi.org/10.2337/DIACLIN.24.2.91
11. Liu, C., You, J., Zhu, W., Chen, Y., Li, S., Zhu, Y., Ji, S., Wang, Y., Li, H., Li, L., Fan, S.: The COVID-19 Outbreak Negatively Affects the Delivery of Care for Patients With Diabetic Foot Ulcers. Diabetes Care. 43, e125–e126 (2020). https://doi.org/10.2337/DC20-1581
12. Kaile, K., Fernandez, C., Godavarty, A.: Development of a Smartphone-Based Optical Device to Measure Hemoglobin Concentration Changes for Remote Monitoring of Wounds. Biosens. 2021, Vol. 11, Page 165. 11, 165 (2021). https://doi.org/10.3390/BIOS11060165
13. Kaile, K., Leiva, K., Fernandez, C., Wu, W., Weigelt, M., Espinosa, A., Kirsner, R., Godavarty, A.: Tissue oxygenation measurements in diabetic foot ulcers using a smartphone based near infrared imaging device. https://doi.org/10.1117/12.2578774. 11632, 1163209 (2021). https://doi.org/10.1117/12.2578774
14. Statistics About Diabetes | ADA, https://www.diabetes.org/resources/statistics/statistics-about-diabetes
15. Fitzpatrick, T.B.: The Validity and Practicality of. Arch. Dermatol. 124, 869–871 (1988)
16. K. Kaile, M. Sobhan, A. Mondal, and A.G.: Machine learning algorithms to classify Fitzpatrick skin types during tissue oxygenation mapping. In: Biophotonics Congress: Biomedical Optics 2022 (Translational, Microscopy, OCT, OTS, BRAIN), Technical Digest Series (Optica Publishing Group, 2022). p. JM3A.4 (2022)

17. Tibshirani, R.: Regression Shrinkage and Selection Via the Lasso. J. R. Stat. Soc. Ser. B. 58, 267–288 (1996). https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
18. Cai, D., Zhang, C., He, X.: Unsupervised feature selection for Multi-Cluster data. Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 333–342 (2010). https://doi.org/10.1145/1835804.1835848
19. Genuer, R., Poggi, J.M., Tuleau-Malot, C.: Variable selection using random forests. Pattern Recognit. Lett. 31, 2225–2236 (2010). https://doi.org/10.1016/j.patrec.2010.03.014
20. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene Selection for Cancer Classification using Support Vector Machines. Mach. Learn. 2002 461. 46, 389–422 (2002). https://doi.org/10.1023/A:1012487302797
21. Yang, Y., Shen, H.T., Ma, Z., Huang, Z., Zhou, X.: L2,1-Norm Regularized Discriminative Feature Selection for Unsupervised Learning.
22. Al Mamun, A., Tanvir, R.B., Sobhan, M., Mathee, K., Narasimhan, G., Holt, G.E., Mondal, A.M.: Multi-run Concrete Autoencoder to Identify Prognostic lncRNAs for 12 Cancers. (2021). https://doi.org/10.1101/2021.08.01.454691
23. Abid, A., Balin, M.F., Zou, J.: Concrete autoencoders: Differentiable feature selection and reconstruction. 36th Int. Conf. Mach. Learn. ICML 2019. 2019-June, 694–711 (2019)
24. Pedregosa FABIANPEDREGOSA, F., Michel, V., Grisel OLIVIERGRISEL, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Dubourg, V., Passos, A., Brucher, M., Perrot andÉdouardand, M., Duchesnay, andÉdouard, Duchesnay EDOUARDDUCHESNAY, Fré.: Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot. J. Mach. Learn. Res. 12, 2825–2830 (2011)
25. Keras: the Python deep learning API, https://keras.io/
26. How does temperature affect softmax in machine learning? | Kasim Te, http://www.kasimte.com/2020/02/14/how-does-temperature-affect-softmax-in-machine-learning.html
27. Sobhan, M., Mamun, A. Al, Tanvir, R.B., Alfonso, M.J., Valle, P., Mondal, A.M.: Deep Learning to Discover Genomic Signatures for Racial Disparity in Lung Cancer. Proc. - 2020 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2020. 2990–2992 (2020). https://doi.org/10.1109/BIBM49941.2020.9313426
28. Van Der Maaten, L., Hinton, G.: Visualizing Data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605 (2008)