Eczema Severity Measurement/Segmentation

Classification

Classification is a type of supervised learning. Classification is a process of categorizing a given set of data into classes. It can be performed on both structured and unstructured data. The process starts with predicting the class of given data points. The classes are often referred to as target, label or categories.

Examples

- K-nearest (KNN)
- Support Vector Machine (SVM)
- Random forest
- Neural networks (CNN)

Clustering

Cluster analysis, or clustering, is an unsupervised machine learning task. Clustering [1] can also be useful as a type of feature engineering, where existing and new examples can be mapped and labelled as belonging to one of the identified clusters in the data.

Some clustering algorithms require the number of clusters to discover in the dataset as input. However, others require the specification of some minimum distance between observations in which examples may be considered "close" or "connected."

2 types of clustering

- **Flat clustering**: In flat clustering, the user specifies the amount of cluster eg: k-means
- **Hierarchical clustering**: The algorithms determines the number of clusters required while the user specifies the min distance eg: mean shift

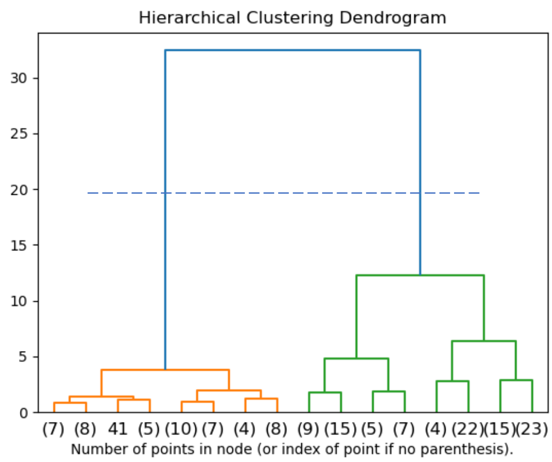

Figure 1 Hierarchical Clustering

Using Neural networks for Clustering

- Self-Organised Maps (SOM)
- Deep Embedded Clustering [2]

Clustering Algorithms [*link*]

Method name	**Parameters**	**Scalability**	**Usecase**	**Geometry (metric used)**
K-Means	number of clusters	*Very large n_samples, medium n_clusters with MiniBatch code*	General-purpose, even cluster size, flat geometry, not too many clusters	Distances between points
Affinity propagation	damping, sample preference	*Not scalable with n_samples*	Many clusters, uneven cluster size, non-flat geometry	Graph distance (e.g. nearest-neighbor graph)
Mean-shift	bandwidth	*Not scalable with n_samples*	Many clusters, uneven cluster size, non-flat geometry	Distances between points
Spectral clustering	number of clusters	*Medium n_samples, small n_clusters*	Few clusters, even cluster size, non-flat geometry	Graph distance (e.g. nearest-neighbor graph)
Ward hierarchical clustering	number of clusters or distance threshold	*Large n_samples and n_clusters*	Many clusters, possibly connectivity constraints	Distances between points
Agglomerative clustering	number of clusters or distance threshold, linkage type, distance	*Large n_samples and n_clusters*	Many clusters, possibly connectivity constraints, non Euclidean distances	Any pairwise distance
DBSCAN	neighborhood size	*Very large n_samples, medium n_clusters*	Non-flat geometry, uneven cluster sizes	Distances between nearest points
OPTICS	minimum cluster membership	*Very large n_samples, large n_clusters*	Non-flat geometry, uneven cluster sizes, variable cluster density	Distances between points
Gaussian mixtures	many	*Not scalable*	Flat geometry, good for density estimation	Mahalanobis distances to centers
Birch	branching factor, threshold, optional global clusterer.	*Large n_clusters and n_samples*	Large dataset, outlier removal, data reduction.	Euclidean distance between points

Proposed approach

- Use clustering algorithms to carry out initial analysis on the obtained dataset
- Study and identify meaningful patterns to ascertain if it can be used to obtain a binary classification of mild or severe eczema.

Research Paper

Automatic Detection and Severity Measurement of Eczema Using Image Processing [3]

The authors have proposed an automatic eczema detection and severity measurement model using modern image processing and computer algorithm. The system can successfully detect regions of eczema and classify the identified region as mild or severe based on image colour and texture feature.

Methodology Used

The authors have used a five-step methodology to solve the problem identified. These steps have been listed as follows.

- **Skin Region Detection**: YCbCr color space has been used to represent skin colours. Additionally, skin images from the web has been used to capture different skin colours under varying lighting conditions. Finally, skin pixel classification is carried out using a Neyman-Pearson hypothesis test [4].
- **Segmentation of Eczema Region**: The authors have used a colour-based segmentation method based on k-means clustering, combined with morphological image processing techniques to segment the affected region.
- **Feature extraction**: This has been done to extract biologically meaningful features of the eczema region that can aid identification and evaluation of the skin disease state. The features that were used include ***colour, texture and border features***. The border feature is the measured as the distance from the centroid to the border of each eczema region. The texture was quantified using a Gray Level Co-occurrence Matrix (GLCM) of a grayscale image.
- **Classification**: The authors used a dataset of labelled images of 3 categories (healthy, mild eczema and severe eczema). Classification have been done using SVM classifier.
- **Severity Index Measurement:** They have used an eczema severity measurement tool EASI [5]. The final severity index is a combination of ***area score, eczema intensity score, and body region score.***

References

[1] J. Brownlee, "10 Clustering Algorithms With Python," *Machine Learning Mastery*, Apr. 05, 2020. https://machinelearningmastery.com/clustering-algorithms-with-python/ (accessed Dec. 15, 2020).

[2] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised Deep Embedding for Clustering Analysis," *arXiv:1511.06335 [cs]*, May 2016, Accessed: Dec. 15, 2020. [Online]. Available: http://arxiv.org/abs/1511.06335.

[3] M. N. Alam, T. T. K. Munia, K. Tavakolian, F. Vasefi, N. MacKinnon, and R. Fazel-Rezai, "Automatic detection and severity measurement of eczema using image processing," in

2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Aug. 2016, pp. 1365–1368, doi: 10.1109/EMBC.2016.7590961.
[4] T. S. Cho, W. T. Freeman, and H. Tsao, “A reliable skin mole localization scheme,” in *2007 IEEE 11th International Conference on Computer Vision*, Oct. 2007, pp. 1–8, doi: 10.1109/ICCV.2007.4409144.
[5] J. M. Hanifin *et al.*, “The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis,” *Experimental Dermatology*, vol. 10, no. 1, pp. 11–18, 2001, doi: https://doi.org/10.1034/j.1600-0625.2001.100102.x.