

REVIEW OF MACHINE LEARNING TECHNIQUES IN OPHTHALMOLOGY: A NOVEL APPROACH

Savita Bamal

E-Mail Id: savitabamal85@gmail.com

School of Engineering and Technology, Ansal University, Gurgaon, India

Abstract-The rising circumstance in this day and age proposes different visual impairment causing maladies like diabetic retinopathy, Age macular degeneration, waterfall, glaucoma. These are significant eye variations from the norm looked in the clinical world. Early analysis and opportune treatment are required to forestall vision misfortune issues. Advanced technology-based computer-aided diagnosis tools using machine learning techniques help to reduce the workload of the ophthalmologist. In this paper, we concentrated on the survey of different existing machine learning models utilized for building up a determination framework for human services applications. This paper features the survey of existing order models alongside their performances.

Keywords: Machine learning, machine learning and healthcare, disease diagnosis using machine learning, classification algorithms, conclusion

1.INTRODUCTION

Machine learning is a subfield of artificial intelligence (AI). The machine learning technologies provide automatic learning ability of system without human intervention. This algorithm focuses on developing Computer-Aided Diagnosis models. These machine learning techniques find a pattern in the dataset that helps in making a better decision support system. Machine learning for healthcare has grown immensely, including work in many domains related to the medical image, computer vision, natural language processing, diagnosing diabetic retinopathy [1], object detection, and large- scale phenotyping from observational data. Various classification-based machine learning techniques make diagnosis system efficient. There are four types of machine learning algorithms.

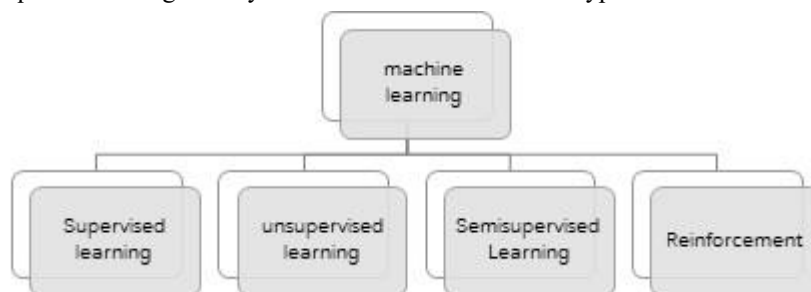


FIG1 MACHINE LEARNING CATEGORY

1.1 Supervised Learning

It comprises of a given arrangement of information factors which are pre-named and target information. Utilizing the information factors, it creates a planning capacity to plan contributions to required yields.

$$Y=f(x) \quad 1$$

The parameter adjustment process continues until the model attains a suitable level of accuracy on the training data. These types of learning algorithms are used in real-time applications. Classification and regression-based problems are example of supervised machine learning.

Classification: This algorithm helps to map input(x) variable to output discrete value(Y). This algorithm predicts results in Yes or No, for example “Is this eye normal”?[2]

Regression: It find correlation between dependent and independent variable. It predicts result on the continuous values. This algorithm helps to map input variable(X) to continuous value(Y).

1.2 Unsupervised learning

In this algorithm, we only have training data rather an outcome data. The system is trained with unlabeled data.

These algorithms are used in pattern detection, to group data points, and in descriptive modeling. It is used in classifiers by recognizing existing patterns or clusters in the input datasets.

Clustering: it creates clusters based on similar groups.

1.3 Semi-Supervised Learning

It comes between supervised and unsupervised learning. This learning use labeled and unlabeled data for training.

1.4 Reinforcement Learning

In this type of algorithm, a machine is trained to map action to a specific decision. It uses various software to find the best output for a given input. The machine trained itself to find the most rewarding actions by reward and punishment using experience [3].

2.MACHINE LEARNING IN HEALTHCARE

Machine learning is a great tool to help humans in various complex situations and systems and that is why these days its highly useful in different applications ranging from aeronautics, real time applications, finance, security, and medical science. Since it has the power of computing and well-written algorithms it can easily do complex stuff and no involvement of humans avoid the risk of manual error which can be life-saving especially in the field of medical science. Machine learning in medical science is normally used to predict diseases by analyzing data from various medical sources and then deriving a pattern from that data. More the data and better its accuracy will result in better prediction. So, data becomes a very important aspect of machine learning.

3. DIAGNOSIS OF THE DISEASE USING MACHINE LEARNING ALGORITHMS

The exact and accurate diagnosis of disease is an important task. The early and exact diagnosis of eye diseases of patients by clinicians can stop blindness problems. Advanced technologies of ML make the task of diagnosis easy and time-saving. Using these effective and efficient machine learning models, there are better results in disease diagnosis problems. This study discussed some of the eye diseases diagnosed by MLT like diabetic retinopathy [3], glaucoma [2], cataract [4], Hyperemia conjunctival [5].

3.1 Cataract Disease Detection

Deepika et al. [3] proposed a method to diagnosis cataract using statistical features and classified with K-means and the Adaptive neural fuzzy interface system (ANFIS). Better accuracy achieved 93% after using this K-means learning model as compared to ANFIS. The limitation faced with the ANFIS model was of not provided better accuracy because of error possibility in the backward pass. In the future, various machine learning algorithms can be used.

Yang et al. [4] proposed an automatic ensemble learning-based for cataract detection. SVM and BPNN were used for two-class classification and the ensemble learning model was used for multiple classes. The accuracy obtained for two classes and multiple classes was 93.2% and 84.5%. It was applicable to single learning models. The limitation was, not give significant accuracy for multiple learning models.

3.2 Hyperemia Conjunctival

Park and Sandberg [6] proposed the methodology of transforming the extracted features into grading scales using a radial basis function network (RBF). The results obtained from the RBF network were variable from feature to feature. Baum [7] implemented the extraction of a region of interest from the best frame. Later various features were computed from the region of interest. Lastly, a multi-layer perceptron (MLP) was used to transform these features into specialists grading scales. The study focused only on a limited set of features for evaluation. Luisa, Sanchez, Noelia Barreira¹, Hugo Pena, and Eva Yebra [8] in 2015 demonstrated an automatic method to measure the hyperemia level of bulbar conjunctiva by transforming extracted features to grading scales (Efron, CCLRU) with an artificial neural network. The suitable frames for grading were selected and ROI (region of interest) was delimited. Then Features were extracted by edge detection algorithm and some morphological operation and transform by ANN.

Mohd. Adnan al. [9] suggested an automated model based on image processing. The advanced technology called

OCULUS Keratography 5M using R scan method was used to accurately diagnosed eye redness. This study focused on the relation of image segmentation as image analysis. Segmentation faced many problems such as noise, poor contrast, color intensity, the limited spatial resolution that make segmentation as a difficult task. But this was the first automated and objective technique to measure bulbar redness.

Huntjens et al. [10] developed an Advanced ophthalmic system (AOS) software to grade conjunctival redness. In this research, it was found that in the objective method, the result obtained by the computing coefficient of repeatability COR (0-10 to 0-13) and then comparing with subjective grading scales (Efron, CCLRU) was consistent. The outcomes achieved through subjective and automated software for conjunctival redness was not interchangeable. The Proposed automatic software provided an instant analysis of digital images using mobile phone applications. But this study had challenges in the standardization of image quality and magnification as the images were taken from various sources.

Yann et al. [11] suggested convolutional network architecture and its application in commercial areas, vision systems, language understanding. This research focused on how to trained convolution networks with gradient-based supervised learning and unsupervised learning algorithms.

Verna et al. [12] developed a CNN (convolutional neural network) based classifier to classify red and white eyes. In this study, it was suggested that no subjective and objective technique used for dependency on operators for evaluation and feature extraction. The model deployed in this study automatically extracts features using MATLAB and then feeding into CNN. It helped in rural areas where patients to clinician's ratio were less. It was observed that the validation accuracy is 93% with validation loss of 22%. This study computed accuracy only for binary images.

3.3 Glaucoma

Glaucoma is the second leading cause of blindness. It deals with a cluster of disease, not to a single disease. Some research in Automatic detection of glaucoma has done by different researchers using various types of machine learning techniques. Different learning models along with their performance were discussed for glaucoma detection.

V. Nicolae et al. [13] proposed a neural network and a fuzzy-based learning model for automated glaucoma detection. Further, R. Mcintyre et al. [14] proposed a decision tree-based algorithm like ID3 for detecting glaucoma with 85% accuracy. A backpropagation neural network-based model was developed using a gaussian filter for preprocessing methods. The accuracy achieved by this model was 84.6%, Khalil et al. [15] suggested an automated method for optical nerve analysis detection for glaucoma. The SVM classification learning model was used for glaucoma disease detection. It achieved an accuracy of 80%. Further, many learning models based on machine learning were discussed. K-NN, naïve Bayes model along with 86% accuracy was presented [16]. a linear regression model with accuracy 99% was used for automated glaucoma disorder in eye images [17]. Rajya guru et al. [2] developed an automated detection system for glaucoma based on feature representation in a random domain. SVM and random forest models were trained using a hybrid feature extractor method.

3.4 Diabetic Retinopathy

It is a serious diabetic eye disease. The presence of high glucose in blood damaged the vessels, those supply important nutrients, and oxygen to the retina. Various studies were discussed for the automated detection of retinal disease. A generalized feed-forward model to diagnose diabetic retinopathy was developed by using retinal images. The accuracy achieved by this model was 94.5% [18]. An automated detection method of Diabetic Retinopathy (DR) was developed to improve accuracy with various feature extraction methods on retinal images. The ensemble classifier was used to classify the DR and non-DR cases. Various classification algorithms like linear regression, support vector machine, the random forest was used to achieve better accuracy (91%) [19].

At last, in this section, diseases diagnosed by Machine Learning Techniques is shown by block diagram represented by Figure2. This block diagram shows various disease diagnosis models along with performance. There are some eye diseases like glaucoma, cataract, diabetic retinopathy, Hyperemia conjunctival along with classification algorithms used in these studies. With the help of a Figurative approach, we can easily compare such

classification algorithms.

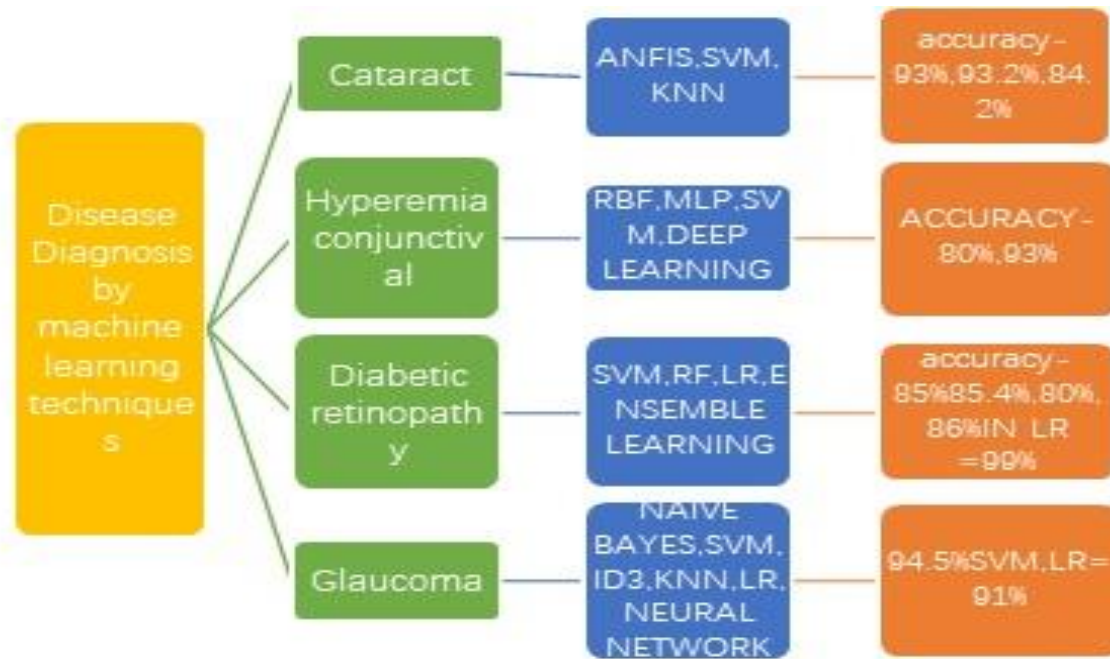


Fig. 3.1 Review of Supervised Technique With Performances For Disease Diagnosis

4. CLASSIFICATION ALGORITHMS

In various machine-learning applications, classification model is used. Data classification is divided into two phases. One is training and the other is classification and its performance is computed using a testing set of data. Based on these algorithms, various classifiers are designed to analyze medical data and helps in disease identification. Some of the classification algorithms are discussed in this section. These are

- Bayesian Based (Naïve Bayes)
- Support vector machine (SVM)
- Artificial Neural Network (ANN)
- K-Nearest Neighbor
- Decision Tree

4.1 Naïve Bayes Classification

It is a probabilistic classifier based on Bayes theorem. It helps in building an efficient learning model for quick prediction. It is simple and makes the classification easy [20]. Based on class labels it identifies the class membership probabilities.

Example for Bayesian classifiers are statistical classifiers.

$$P(A|B) = \frac{p(B|A)p(A)}{P(B)} \tag{2}$$

Where A and B are events and we are trying to find out the probability of A.

Here P(A)= prior probability, P(A|B) is a posteriori probability of B.

4.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) was designed in the 1990s. SVM is an important process for the classification model. To accomplish ML errands, a help vector machine (SVM) is utilized, and it is a straightforward and unmistakable procedure. During this strategy, an assortment of preparing tests is given, each example is isolated into various classifications. It performs well on high dimensional data and avoids the curse of dimensionality problem [21]. SVM is trained with a dataset to perform the task of prediction and classification. Here linear SVM is defined as

$$f(v) = W^T X + b \tag{3}$$

where Xi is data point from the sample set, W is weight vector and b is bias

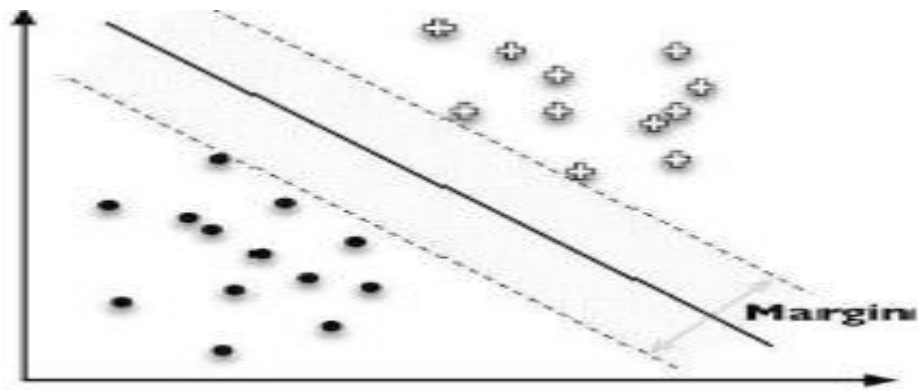


Fig. 4.1 Maximum Margin Hyperplane SVM Model [21]

4.3 Decision Tree

Choice Tree/Decision Tree (DT) is a regulated technique utilized for both grouping and relapse (CART). It is a prescient methodology. It is a tree-like structure where the inside hub used to test highlights and leaf hub are spoken to by class names. The way from the root to the leaf is taken characterization rule. To settle on the choice tree, idea of data gain (IG) is used. This is famous on the grounds that the development is straightforward, which is non-parametric [22].

4.4 Neural Network

This network is known as an artificial neural network and was invented by Dr. Robert. This network can be defined as a computing system made-up of the number of processing units called neurons. This network has one input layer, one hidden layer, one hidden layer, and one output layer. Feedforward and backward propagation are used to train the models [23].

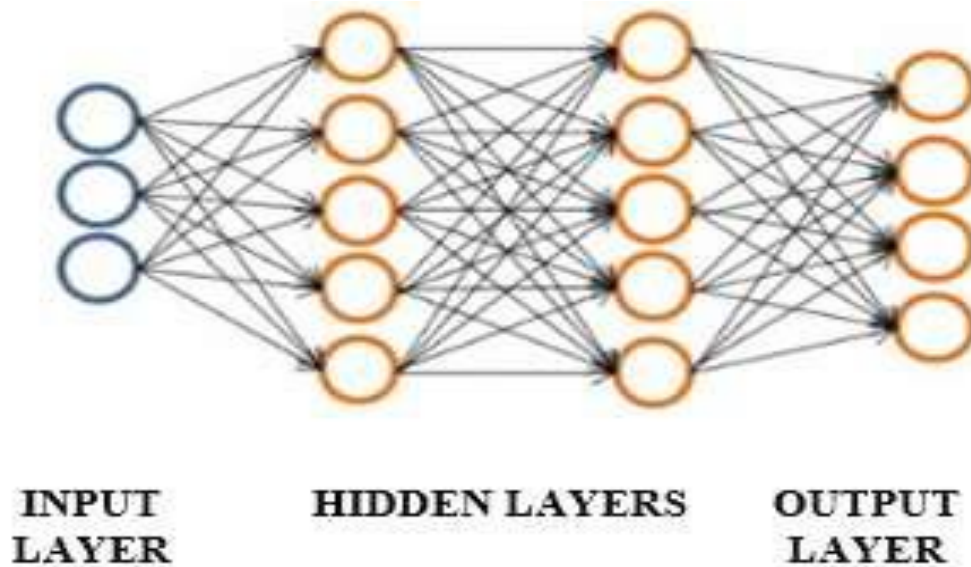


Fig. 4.2 Artificial Neural Network Model [23]

4.5 K-Nearest Neighbor (KNN)

K-Nearest neighbor calculation (KNN) is supervised learning calculations that have been utilized in numerous applications in the field of information mining, factual example acknowledgment, and numerous others. It follows a technique for characterizing objects dependent on nearest preparing models in the component space. K-closest neighbor is an every now and again utilized methodology for the order of tests. With this procedure, we can compute the separation measure from the N number of the preparation test.

The performance measures like Accuracy, Precision, Recall, F-score can be evaluated by the help of the evaluation matrix. This matrix is also known as a confusion matrix. It has actual observations and predicted observations shown in the below matrix

Table-4.1 Evaluation Matrix for Classifier Performance

Confusion matrix		Predicted class	
		Yes	No
Actual Class	Yes	TP	FN
	No	FP	TN

On the bases of performance measures, different researchers have computed accuracy of all learning models. These performance measures are defined as given below equations[24].

$$\text{Accuracy} = (TP + TN) / (TP + FP + TN + FN) \tag{4}$$

Recall- It measures the fraction of positive observations correctly predicted by classifier.

$$\text{Recall} = TP / (TP + FN) \tag{5}$$

Precision- It is ratio of correctly predicted positive observations to the total predicted positive observations. The higher the precision is, the lower the number of false positive error rate. It is defined as

$$\text{Precision} = TP / (TP + FP) \tag{6}$$

CONCLUSION

Supervised ML based models are summed up in this survey paper. Different models have utilized in the field of ophthalmology alongside their performances. From past investigations, it has been discovered that the precision of Linear regression and SVM classifier is acceptable in some infection finding models. The presence of computer-aided diagnosis tools based on machine learning techniques makes ophthalmologist work simple and productive. These days, a great deal of work is going on tackled clinical world issues. In this investigation, specialists are attempting to improve the presentation of the model. This is a very challenging task for specialists to choose fitting calculations in light of the fact that each calculation execution is diverse because of various info datasets. Later on, different calculations have been intended for the early conclusion of eye issues. AI is an open region for research that will give better procedures to plan a productive model for illness analysis. Later on, we will attempt to improve the precision of model for analysis of diabetic retinopathy utilizing grouping calculations.

REFERENCES

- [1] A. Bhatkar, "GFF classifier for detection of diabetic Retinopathy in retinal images," *Int. J. Comput. Appl.*, vol. 134, no. 14, 2016.
- [2] V. Rajyaguru, C. Vithalani, and R. Thanki, "ORIGINAL RESEARCH A literature review : various learning techniques and its applications for eye disease identification using retinal images," *Int. J. Inf. Technol.*, 2020.
- [3] E. deepika and S. Maheswari, "Detection of cataract by statical features and classification," *Int. Res. J. Eng. Technol.*, vol. 5, no. 3, 2018.
- [4] J. Yang et al., "Acceptance test," *Comput. Methods Programs Biomed.*, 2015.
- [5] P. Fieguth and T. Simpson, "Automated measurement of bulbar redness," *Investig. Ophthalmol. Vis. Sci.*, vol. 43, no. 2, pp. 340–347, 2002.
- [6] J. Park and I. Sandberg, "Approximation and Radial-Basis-Function Networks," *Neural Comput.*, vol. 5, no. 2, pp. 305–316, 1993.
- [7] E. Baum, "On the capabilities of multilayer perceptrons," *J. Complex.*, vol. 4, no. 3, pp. 193–215, 1988.
- [8] L. Sanchez, N. Barreira, H. Pena-Verdeal, and Y.-P. Eva, "A Novel Framework for Hyperemia Grading Based on Artificial Neural Networks," in *International Work-Conference on Artificial Neural Networks*, 2015, vol. 9094, pp. 5–7.
- [9] M. R. H. Mohd Adnan, A. Mohd Zain, H. Haron, R. Alwee, M. Zulfaezal Che Azemin, and A. Osman Ibrahim, "Eye Redness Image Processing Techniques," *J. Phys. Conf. Ser.*, vol. 892, no. 1, 2017.
- [10] B. Huntjens, M. Basi, and M. Nagra, "Contact Lens and Anterior Eye Evaluating a new objective grading

- software for conjunctival hyperaemia,” *Contact Lens Anterior Eye*, no. July, pp. 1–7, 2019.
- [11] Y. LeCun and C. Kavukcuoglu, Koray Farabet, “Convolutional networks and applications in vision,” in *IEEE International Symposium on Circuits and Systems*, 2010.
- [12] S. Verma, L. Singh, and M. Chaudhry, “Classifying Red and Healthy Eyes using Deep Learning,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 7, pp. 525–531, 2019.
- [13] V. Nicolae, “Computational intelligence for medical knowledge acquisition with application to glaucoma,” in *First IEEE International Conference on Cognitive Informatics*, 2002.
- [14] A. R. McIntyre, “oward glaucoma classification with moment methods,” in *First Canadian Conference on Computer and Robot Vision*, 2004. *Proceedings*, 2004.
- [15] T. Khalil, S. Khalid, and A. M. Syed, “Review of Machine Learning techniques for glaucoma detection and prediction,” *Proc. 2014 Sci. Inf. Conf. SAI 2014*, pp. 438–442, 2014.
- [16] A. Dey and S. K. Bandyopadhyay, “Automated Glaucoma Detection Using Support Vector Machine Classification Method,” vol. 11, no. 12, pp. 1–12, 2016.
- [17] Y. Hagiwara et al., “Computer-aided diagnosis of glaucoma using fundus images: A review,” *Comput. Methods Programs Biomed.*, vol. 165, pp. 1–12, 2018.
- [18] Sudhir. S. Kanade, “An Amalgamation-Based System for Micro aneurysm Detection and Diabetic Retinopathy Grading,” *Ijmer*, vol. 5, no. 1, pp. 70–77, 2015.
- [19] K. Bhatia, S. Arora, and R. Tomar, “Diagnosis of diabetic retinopathy using machine learning classification algorithm,” *Proc. 2016 2nd Int. Conf. Next Gener. Comput. Technol. NGCT 2016*, no. October, pp. 347–351, 2017.
- [20] P. R. Harper, “A review and comparison of classification algorithms for medical decision making,” *Health Policy (New York)*, vol. 71, no. 3, pp. 315–331, 2005.
- [21] P. S. Helode, Dr. K. H. Walse, and Karande M.U., “An Online Secure Social Networking with Friend Discovery System,” *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 5, no. 4, pp. 8198–8205, 2017.
- [22] A. M. Almana, S. Aksoy, and R. Alzahrán, “A Survey On Data Mining Techniques In Customer Churn Analysis For Telecom Industry,” *J. Eng. Res. Appl. www.ijera.com*, vol. 4, no. 5, pp. 165–171, 2014.
- [23] Y. Tong, W. Lu, Y. Yu, and Y. Shen, “Application of machine learning in ophthalmic imaging modalities,” *Eye Vis.*, vol. 7, no. 1, pp. 1–15, 2020.
- [24] L. S. Brea, N. Barreira, A. Mosquera, H. Pena-Verdeal, and E. Yebra-Pimentel, “Comparing machine learning techniques in a hyperemia grading framework,” *ICAART 2016 - Proc. 8th Int. Conf. Agents Artif. Intell.*, vol. 2, no. Icaart, pp. 423–429, 2016.
- [25] V. K. Pang-Ning Tan, Michael Steinbach, *Introduction to Data Mining*. 2016.