A Fuzzy based technique for Pattern Recognition & Classification

M. D. Anto Praveena*, A.Christy, L.Suji Helen, S. Jancy, D. UshaNandini
Department of Computer Science & Engineering, Sathyabama Institute of Science & Technology, Chennai, India – 600119.

antopraveena@gmail.com

Abstract. A pattern can either be seen genuinely or it tends to be watched numerically by applying calculations. Pattern classification is worried about the capacity to discover absolute names for a lot of perceptions. A pattern classification task was viewed as an example determination issue where a meagre subset of test from the marked preparing set was picked. We proposed a versatile learning calculation using the least square capacity to address this issue. Utilizing these chose tests, which we call educational vectors, a classifier equipped for perceiving the test tests was built up. This epic calculation is a mix of looking through systems that, in light of forward looking through advances, yet Adaptive finds a way to address the blunders presented by before forward advances. This paper reviews cost-delicate fuzzy standard based frameworks for pattern classification. Weighted preparing patterns are utilized to build cost-touchy fuzzy principle-based frameworks. A fuzzy classification framework is built from a given arrangement of preparing patterns. It is accepted that a weight is appointed to each preparation pattern from the earlier. The heaviness of preparing patterns can be determined dependent on their dispersion.

1. Introduction

Pattern acknowledgment is the logical order whose design is to group patterns (otherwise called occasions, tuples and models) into a lot of classifications which are likewise alluded to as classes or marks. Generally, the classification depends on factual models that are actuated from an excellent arrangement of pre classified patterns. On the other hand, the classification uses information that is provided by a specialist in the application space. A pattern is generally made of a lot of estimations that portray a specific item. Another regular application which utilizes pattern acknowledgment is Optical character acknowledgment. These applications convert filtered reports into machine-editable content to improve their capacity and recovery. Each record experiences three stages. Initial, an administrator checks the record. This proselytes the record into a bitmap picture. Next, the examined report is fragmented with the end goal that each character is detached from the others. At that point, a component extractor gauges certain highlights of each character, for example, open zones, shut shapes, corner to corner lines and line crossing points. At last, the filtered characters are related to their comparing alpha-numeric character [1]. The affiliation is gotten by applying a pattern acknowledgment calculation to the highlights of the checked characters Fig [1]. Right now, set of names/classifications/classes are the arrangement of alpha-numeric character for example letters, numbers, accentuation marks, and so on.

In a commonplace factual pattern acknowledgment setting, a lot of patterns S, additionally
alluded to as a preparation set is given. The marks of the patterns in S are known and the objective is to develop a calculation so as to name new patterns. A classification calculation is otherwise called an induce rand a case of an inducer for a particular preparing set is known as a classifier.

The preparation set can be depicted in a assortment of ways. Most as often as possible, each pattern is portrayed by a vector of highlight values. Every vector has a place with a solitary class and connected with the class name. Right now, preparing set is put away in a table where each column comprises of an alternate pattern. An acceptance calculation, or all the more compactly an inducer (otherwise called student), is a calculation that is given a preparation set and builds a model that sums up the association between the information qualities and the objective trait. For instance, an inducer may take as info explicit preparing patterns with their relating class names, and produce a classifier [2].

$$E(I(S), D) = \sum_{(x,y) \in U} D(x,y).L(y, I(S)(x))$$

Rule enlistment calculations create a lot of in the event that decides that depicts the classification procedure. The principle bit of leeway of this methodology is its high intelligibility. To be specific the guidelines can be composed as an assortment of continuous contingent proclamations in plain English which hare anything but difficult to utilize. A large portion of the Rule acceptance calculations depend on the different and vanquish worldview.

![FIG 1 CLASSIFIER](image)

In building up a classifier for a given issue we continue determining input factors, specifically includes, gathering input–yield sets as indicated by the decided highlights, preparing the classifier and assessing classifier execution. In preparing the classifier, extraordinary consideration must be taken with the goal that test the dataset is utilized for preparing the classifier. Assume that a classifier for an n class issue is tried utilizing M data tests [3]. To assess the classifier for a test data set, we produce an n × n perplexity lattice A, whose component aim is the quantity of class I data grouped into class j. At that point the acknowledgment rate R or acknowledgment exactness in % is determined

2. Core Concepts

2.1 Check Data Quality

Hope to spend a decent part of your undertaking time doing data adjustment and purging, now and again up to 80%. The recovery of data is the first occasion when you’ll review the data in the data science process. The greater part of the mistakes you’ll experience during the data gathering stage are anything but difficult to spot, yet being too reckless will cause you to spend numerous hours comprehending data gives that could have been forestalled during data import. During data recovery, you verify whether the data is equivalent to the data in the source record and hope to check whether you have the correct data types. This shouldn't take excessively long; when you have enough proof
that the data is like the data you find in the source report, you stop. With data planning, you do a progressively intricate check. In the event that you worked superbly during the past stage, the mistakes you discover now are additionally present in the source archive [4]. The emphasis is on the substance of the factors: you need to dispose of grammatical errors and other data section mistakes and carry the data to a typical standard among the data sets. For instance, you may address USQ to USA and United Kingdom to UK. During the exploratory stage your center movements to what you can gain from the data. Presently you expect the data to be perfect and take a gander at the factual properties, for example, dispersions, relationships, and anomalies. You'll regularly repeat over these stages [5].

2.2 Cleansing, Integrating, and Transforming data

The data got from the data recovery stage is probably going to be "a gem waiting to be discovered." Your assignment currently is to purify and set it up for use in the demonstrating and announcing stage. Doing so is immensely significant in light of the fact that your models will perform better and you'll lose less time attempting to fix weird yield [6]. It can't be referenced model needs the data in a particular organization, so data change will consistently become an integral factor. It's a decent propensity to address data mistakes as at an early stage in the process as would be prudent. Be that as it may, this isn't constantly conceivable in a sensible setting, so you'll have to take restorative activities in your program.

\[
L(y, I(S)(x)) = \begin{cases} 
0 & \text{if } y = I(S)(x) \\
1 & \text{if } y \neq I(S)(x) 
\end{cases}
\]

2.3 Data Cleansing

Data purifying is a sub process of the data science process that centers around evacuating mistakes in your data so your data turns into a genuine and steady portrayal of the procedures it begins from. By genuine and steady portrayal" we suggest that at any rate two kinds of mistakes exist. The primary kind is the translation mistake, for example, when you underestimate the incentive in your data for without a doubt such as saying that an individual's age is more noteworthy than 300 years. These kind of blunder focuses to irregularities between data sources or against your organization's institutionalized qualities. A case of this class of mistakes is putting "Female" in one table and "F" in another when they speak to something very similar: that the individual is female. Another model is that you use Pounds in a single table and Dollars in another.

2.4 Data Entry Errors

Data assortment and data section are blunder inclined procedures. They regularly require human intercession, and on the grounds that people are just human, they make grammatical errors or lose their focus for a second and bring a mistake into the chain. Be that as it may, data gathered by machines or PCs isn't liberated from mistakes either. Instances of blunders beginning from machines are transmission mistakes or bugs in the concentrate, change, and burden stage (ETL) [7]. For little data sets you can check each an incentive by hand. Distinguishing data mistakes when the factors you study don't have numerous classes should be possible by arranging the data with tallies. At the point when you have a variable that can take just two qualities: "Great" and "Awful", you can check whether those are genuinely the main two qualities present.

2.5. Outliers

An outlier is a perception that is by all accounts far off from different perceptions or, all the more explicitly, one perception that follows an unexpected rationale or generative procedure in comparison to different perceptions. The least demanding approach to discover outliers is to utilize a plot or a table with the base and greatest qualities [8].
The plot on the top shows no outliers, though the plot on the base show’s potential outliers on the upper side when a typical dissemination is normal. The typical circulation, or Gaussian dissemination, is the most well-known appropriation in characteristic sciences. It shows most cases happening around the normal of the dispersion and the events diminish when further away from it. The high qualities in the base diagram can highlight outliers while accepting an ordinary circulation. As we saw before with the relapse model, outliers can gravely impact your data demonstrating, so explore them first.

2.6 Dealing with Missing Values

Missing values aren’t really off-base, yet you despite everything need to deal with them independently; certain displaying systems can't deal with missing values. They may be a marker that something turned out badly in your data assortment or that a mistake occurred in the ETL procedure. Which procedure to use at what time is reliant on your specific case [9]. On the off chance that, for example, you don't have perceptions to save, precluding a perception is most likely impossible.

3. Proposed system

Fuzzy min–max (FMM) is a special type of neural network joins the activities of counterfeit neural network and fuzzy set hypothesis is into a typical structure. FMM is viewed as one of the most valuable neural networks for pattern classification. The proposed framework performs
- Investigate the FMM neural network as far as its effect in tending to pattern classification issues.
- Examine models that are proposed dependent on the first FMM model (i.e., existing FMM-based variations).
- Recognize the difficulties related with FMM and its variations.
- Talk about future patterns and make proposals for development.

The proposed framework gives more precision less number of highlights and furthermore with less number of hyper boxes. All highlights are not important to give classification choice, exact classification should be possible with minimum number of highlights. The computational and network intricacy is additionally diminished. The proposed framework arranges input patterns with less no. of hyper box made and better exactness accomplished at the min estimation of territory of hyper box than size of hyper box merged with models. Henceforth with this calculation we get less misclassification rate, and less perplexing structure at min estimation of zone of hyper box, with min preparing tests.

FMM is equipped for building a nonlinear choice limit of any shape to isolate data tests from various objective classes. Then again, the objective classes will cover with each other, consequently making the issue of covering classes. FMM can construct a nonlinear choice limit to minimize the level of misclassification by eliminating covering areas of various classes. In addition, when compared with other ANN models FMM will take short time to prepare, for example, back propagation, course relationship and Boltzmann machine. This is because of the learning calculation of FMM that requires just a solitary go without emphasis through the data tests.

Advantages
- Improve the FMM performance for solving pattern classification problems.
- Produces a more accurate decision boundaries
- Increases the accuracy rate
- Reduces Network Complexity
- Improve Recall time
3.1 Architecture

![Architecture Diagram]

Fig 2 Architecture

4. Results and Discussion

![Average Price Yearly Line Plot]

Fig 3 Line Plot For Average Price Yearly

After performing the data cleaning techniques on the datasets, calculating average price for two types of the data. The above fig 3 shows the line plot for average price for every year.

![Total Volume Point Plot]

Fig 4 Point Plot For Total Volume Based On Every Month

In the above fig 4, shows the point plot for total volume for every month of an year after calculating the total volume of the data.

![Average Price Region Line Graph]

Fig 5 Line Graph For Average Price Based On Region
In the above fig 5, shows the line plot for average price for every year in every region for future prediction of the usage.

![Fig 5 Pattern Classification For Train And Test Dataset](image)

Above fig 6, is the final output shows the scatter plot for pattern classification for train and test dataset.

5. Conclusion

The proposed framework furnishes more exactness with less number of highlights and furthermore with less number of hyper boxes. All highlights are not important to give classification choice, precise classification should be possible with minimum number of highlights. The computational and network intricacy is additionally diminished. The proposed framework groups input patterns with less no. of hyper box made and better precision accomplished at the min estimation of territory of hyper box than size of hyper box merged with models. Consequently, with this calculation we get less misclassification rate, and less mind-boggling structure at min estimation of territory of hyper box, with min preparing tests.

References

[1] Muslima, Umme, and M. Babul Islam. "Experimental framework for mel-scaled LP based Bangla speech recognition." Computer and Information Technology (ICCIT), 2013 16th International Conference on. IEEE, 2014.

[2] Paul, Anup Kumar, Dipankar Das, and Md Mustafa Kamal. "Bangla speech recognition system using LPC and ANN." Advances in Pattern Recognition, 2009. ICAPR'09. Seventh International Conference on. IEEE, 2009.

[3] Christy, A., Vaithyasubramanian, S., Mary, V.A., Naveen Renold, J. (2019)," Artificial intelligence based automatic decelerating vehicle control system to avoid misfortunes ",International Journal of Advanced Trends in Computer Science and Engineering,Vol.8,Issue.6, Pp.3129-3134.

[4] Budiharto, Widodo, and DerwinSuhartono. "Intelligent service robot with speech recognition and telepresence capabilities." In SAI Intelligent Systems Conference (IntelliSys), 2015, pp. 301-304. IEEE, 2015.

[5] Haro, Luis Fernando D., "Low-cost speaker and language recognition systems running on a Raspberry Pi." IEEE Latin America Transactions 12.4 (2014): 755-763.

[6] Thiang, DhannyWijaya. "Limited speech recognition for controlling movement of mobile robot implemented on atmega162 microcontroller." In Computer and Automation Engineering, 2009. ICCAE'09. International Conference on, pp. 347-350. IEEE, 2009.

[7] Gandhi M,ReddyK,Hemanth K, “Gis enabled black box system for accident alerts”, International Journal of Engineering and Advanced Technology (2018), ISSN: 2249 – 8958, Volume-7 Issue-4, pp.100-102.

[8] Aravind, K.R.N.V.V.D., PraylaShyry, S., Felix, Y.,” Classification of Healthy and Rot Leaves of Apple Using Gradient Boosting and Support Vector Classifier”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-12, October 2019.
[9] Fezari, Mohamed. "New speech processor and ultrasonic sensors based embedded system to improve the control of a motorised wheelchair." Journal of Applied Sciences Research5.10 (2009): 1750-1755.
[10] Fezari, Mohamed, and Mounir Bousbia-Salah. "A speech command system for autonomous robots guidance." In Advanced Motion Control, 2006. 9th IEEE International Workshop on, pp. 261-265. IEEE, 2006.
[11] Sajkowski, M. "Speech control of dual-drive mobile robots-survey of algorithms." In Robot Motion and Control, 2002. RoMoCo'02. Proceedings of the Third International Workshop on, pp. 387-392. IEEE, 2002.
[12] Kubik, T., and M. Sugisaka. "Use of a cellular phone in mobile robot speech control." In SICE 2001. Proceedings of the 40th SICE Annual Conference. International Session Papers, pp. 106-111. IEEE, 2001.
[13] Pleshkova, Snejana, and Zahari Zahariev. "Development of system model for audio visual control of mobile robots with speech and gesture commands." In Electronics Technology (ISSE), 2017 40th International Spring Seminar on, pp. 1-4. IEEE, 2017.
[14] G. Sai Kirthiga and S. Vaithyasubramaniam (2016). SURVEY on secured online authentication and defence technique against third party human attacks using CAPTCHA. Journal of Chemical and Pharmaceutical Sciences, 9(4), Pp. 2009-2011.
[15] G. Sai Kirthiga and S. Vaithyasubramaniam (2016). Review on development of some strong visual CAPTCHAs and breaking of weak audio CAPTCHAs. In 2016 International Conference on Information Communication and Embedded Systems (ICICES), Pp. 1-4. IEEE.
[16] Wang, Duojin, and Hongliu Yu. "Development of the control system of a speech-operated wheelchair with multi-posture characteristics." In Intelligent Robot Systems (ACIRS), 2017 2nd Asia-Pacific Conference on, pp. 151-155. IEEE, 2017.
[17] Mishra, Anurag, Pooja Makula, Akshay Kumar, Krit Karan, and V. K. Mittal. "A speech-controlled personal assistant robot." In Industrial Instrumentation and Control (ICIC), 2015 International Conference on, pp. 523-528. IEEE, 2015.
[18] Xiaoling Lv and Minglu Zhang and Hui Li. Robot Control Based on Speech Command. 2008 IEEE International Conference on Automation and Logistics Year: 2008 Pages: 2490 – 2494.
[19] Fezari, Mohamed. "New speech processor and ultrasonic sensors based embedded system to improve the control of a motorised wheelchair." Journal of Applied Sciences Research5.10 (2009): 1750-1755.
[20] Imteaj, A., Rahman, T., Hossain, M.K., Alam, M.S. and Rahat, S.A., 2017, February. An IoT based fire alarming and authentication system for workhouse using Raspberry Pi 3. In Electrical, Computer and Communication Engineering (ECCE), International Conference on (pp. 899-904). IEEE.
[21] Imteaj, Ahmed, Tanveer Rahman, Muhammad Kamrul Hossain, and Satka Zaman. "IoT based autonomous percipient irrigation system using raspberry Pi." In Computer and Information Technology (ICCIT), 2016 19th International Conference on, pp. 563-568. IEEE, 2016.
[22] Jesudoss A., R. Vybhavi and B. Anusha, “Design of Smart Helmet for Accident Avoidance”, IEEE International Conference on Communications and Signal Processing, organized by Adhiparasakthi Engineering College, Chennai, April 4 to 6, 2019.
[23] Indhuja, D., Dharsanya, R., Anto Praveena, M.D.,” Auto precautionary alert and accident avoidance system using image processing” pp.151-156, 2016.