A Review on Quantum Clustering in Document Clustering

Bharat Shajhlan
Vaish College of Engineering, Rohtak, Haryana
bharat9214@gmail.com

Abstract

This paper discuss about quantum clustering. Quantum clustering is the technique of data classification tasks. Quantum clustering is a technique which derived from quantum mechanics. The usage of Schrödinger equation in clustering termed as Quantum clustering.
Keywords: Hilbert Space, Schrödinger Equation, Parzen Window Estimator, Quantum Clustering.

1. Introduction

As growing amount of data, data is generated in huge amount via various companies, institutions, organisations. So various ideas came into existence for classifying data, such type of classification is known as Big Data Classification. For Big data classification, clustering techniques were used. One of the clustering techniques is Quantum clustering.

The original algorithm suggested by Horn and Gottlieb, which lead to introduce quantum mechanics in clustering techniques. There are many terms in Quantum Mechanics. These terms are the basics to understand Quantum clustering such as Gaussian kernel, Hilbert Space, Schrödinger equation. Some other terms are Probability distribution, normalization etc. Now, one-by-one these terms is explained as follows:

1.1 Gaussian Kernel

The Gaussian (better Gaußian) kernel is named after Carl Friedrich Gauß (1777-1855), a brilliant German mathematician. The Gaussian kernel is defined in 1-D, 2D and N-D respectively as

$$f(x) = ae^{\frac{(x-b)^2}{2c^2}}$$

1.2 Hilbert Space

This concept was named after David Hilbert. It studies the generalization of Euclidean Space. In this the method of vector algebra and calculus from the two-dimensional Euclidean plane and three-dimensional space to spaces with any finite or infinite number of dimensions. A Hilbert space is an abstract vector space possessing the structure of an inner product that allows length and angle to be measured.

1.3 Schrödinger Equation

Schrödinger equation is a partial differential equation that describes how the quantum state of a quantum system changes with time. It is the combination of both potential and kinetic equation of a wave.

$$H\frac{\partial \psi}{\partial t} = -\frac{\sigma^2}{2}\nabla^2\psi + V\psi$$

1.4 Clustering

Clustering is treated as an unsupervised learning technique in machine learning. It's main goal to check out the structures. Structures are hidden in a wide pool of words. Such structures come to be known as Clusters. Clusters can find out naturally. Clusters can be found in dimensional forms. So dimensional representation play a key role in identify a cluster. The number and size varied from problem to problem. There are various techniques used to check out clusters. One of such techniques is Quantum clustering.

2. Document Clustering Using Quantum Clustering

2.1 Quantum Clustering

As in [1],[7] the quantum clustering is defined , in QC a new way of clustering is done. The QC is inspired from quantum mechanics. Like in quantum mechanics, Schrödinger equation is used to find the behaviour of a particle. Similarly, quantum clustering is used to find out the cluster in one or more than one dimensional. In the Quantum clustering, a solution is generated with the help of Schrödinger equation. The solution is generated with the help of approaches of Scale Space algorithm and Support Vector algorithms. In the Scale Space algorithm the probability distribution of a data set is measured by parzen-window estimator and represented in Euclidean space .While in the Support Vector clustering data clustering is represented in Hilbert Space. The Schrödinger equation can be represented as

$$H\psi(x) = (-\frac{\sigma^2}{2}\nabla^2 + V)\psi(x)$$

2.2 Analysing Document using QC

As per[9], Data present in the document are pre-processed & represented as document term matrix (DTM) . This DTM is used by single value decomposition (svd) to reduce the number of features. After the process of svd, resultant matrix will be distributed in Hilbert Space using Schrödinger equation. The computed evolution system is generated. Then computed local minima are computed to find clusters. Further results were analysed.

3. Applications of Quantum Clustering

Quantum Clustering has been used in generating results from various problems. Some applications of Quantum Clustering are as follows:

In Feature Subset Selection for Mammographic Image Selection: This procedure is done in 2 way process, firstly partitioning the feature using Quantum clustering and secondly a selection of a representative for each cluster is carried out. In this way feature selection is done via quantum clustering [3].

In Solving Pattern Recognition Problems: Quantum clustering technique plays a crucial role in solving pattern recognition problems. As in [7], concept of solving pattern is derived from the quantum clustering only. So the fundamentals of solution can be derived with the help of Schrödinger equation of quantum mechanics .The pattern were formed in a problem and after that pattern can be analysed with help the Schrödinger equation. After result can be generated and formed. In this way Quantum clustering solves the Pattern Recognition Problems.

In Text Analysis: Text Analysis can be done clustering .So the idea of text clustering is occurred. Text clustering is a process of document organisation. It is used in text mining. As in [8], Quantum clustering technique is used in text analysis field. In this a document is analysed with the help of quantum clustering. It has been shown that quantum clustering giving a modern machine learning technology. So this gives a fresh approach which can be used in Data analysis.

4. Conclusion

As discussed earlier, that the quantum clustering is used in various application point of view. Such as in image analysis, feature selection, textual analysis and many others. With the help of quantum clustering many problems were solved effectively as compared with other clustering algorithms such as State Space clustering and Vector Space clustering. There is scope in many real life problems using Quantum Clustering.

References

[1] David Horn and Assaf Gottlieb, the Method OF Quantum Clustering School of Physics and Astronomy Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel .

[2] Shreya Jain, Samta Gajbhiye, A Comparative Performance Analysis of Clustering Algorithms ISSN: 2277 128X, 2012.

International Journal of Engineering, Management, Humanities and Social Sciences Paradigms (IJEMHS)
Volume 30, Issue 01, Quarter 01 (January to March 2018)
An Indexed and Referred Journal with Impact Factor: 2.75
ISSN (Online): 2347-601X
www.ijemhs.com

[3] N. Hamdi, K. Auhmani, M.M. Hassani. Quantum Clustering Based Feature Subset Selection for Mammographic Image Selection, 2015.

[4] Nezha Hamdi, Khalid Auhmani, Moha. M'Rabet Hassani. A new approach Based on Quantum Clustering and Wavelet Transform for breast cancer Classification, ISSN: 2088-8708, 2015.

[5] Esma Aimeur, Gilles Brassard, Sebastien Gambs, Quantum Clustering Algorithms, 2007.

[6] The Gaussian Kernel, Mathematics, 1967.

[7] David Horn and Assaf Gottlieb, Algorithm for Data Clustering in Pattern Recognition Problems Based on Quantum Mechanics. Phys. Rev. Lett. 88 (2002) 018702.

[8] Ding Liu, Minghu Jiang, Xiao fang Yang. Quantum Clustering -A Novel Method for Text Analysis. Phys. (2014).

[9] Sahinur Rahman Laskar, Bhagban Swain. Analysis of Document Clustering using Pseudo Dynamic Quantum Clustering Approach (2016)