



# دانشگاه علوم پزشکی کرمان

## دانشکده بهداشت



پایان نامه مقطع کارشناسی ارشد رشته آمار زیستی

عنوان:

اجرای روش خوشه بندی فازی C-means جهت بر آورد تشخیص بیماری های کبدی و مقایسه

آن با روش خوشه بندی کلاسیک

توسط: مهدیه بذرافشان

استاد راهنما: دکتر عباس بهرامپور

استاد مشاور: دکتر محمد جواد زاهدی

سال تحصیلی: ۱۳۹۳-۱۳۹۴

# **Fuzzy Cluster Means System for the prediction of Diagnosing Liver Disease in Comparing with Classic Clustering**

A Thesis  
Presented to  
The Graduate Studies

By

**Mahdieh Bazrafshan**

In Partial Fulfillment  
of the Requirements for the Degree  
Masters

**Biostatistics**

**Kerman University of Medical Sciences**  
**July ۲۰۱۴**



## چکیده

**مقدمه:** کبد بزرگترین غده و پیچیده ترین عضو بدن محسوب می شود که بیش از ۵۰۰ عمل مختلف را انجام می دهد. متخصصین جهت تشخیص بیماری های کبدی با تعداد زیادی از تست های آزمایشگاهی و انبوهی از داده ها روبرو هستند. این پژوهش با استفاده از الگوریتم خوشه بندی فازی در تلاش است که به متخصصین در افتراق و تشخیص بیماری های کبدی شامل سیروز کبدی، سرطان کبد، هپاتیت حاد و مزمن و کبد چرب کمک کند و در پایان نتایج حاصل از آن با خوشه بندی کلاسیک مقایسه گردیده است.

**روش:** در این مطالعه ۱۱۲۳ بیمار کبدی که جهت ثبت متغیرات آنها از چک لیست استفاده گردید با روش تحلیل خوشه ای فازی و کلاسیک مورد بررسی و تجزیه و تحلیل قرار گرفته و گروه بندی شدند.

**یافته ها:** جهت اجرای روش خوشه بندی فازی یکی از معروفترین و پرکاربردترین الگوریتم های خوشه بندی یعنی FCM<sup>۱</sup> مورد استفاده قرار گرفت که پس از اجرای این روش افتراق بین بیماری ها به خوبی صورت گرفت به گونه ای که ۱۰۰ درصد بیماران کبد چرب در خوشه اول، ۶۳ درصد بیماران سرطانی در خوشه دوم، ۸۷ درصد بیماران هپاتیت مزمن در خوشه سوم، ۷۴ درصد بیماران هپاتیت حاد در خوشه چهارم و ۸۴ درصد بیماران سیروز کبدی در خوشه پنجم قرار گرفتند. مطالعه حاضر نشان داد که گرچه اجرای تحلیل خوشه ای فازی و کلاسیک منجر به افتراق تقریباً یکسان بیماریها می گردد و درصد افتراق بیماریهای کبد چرب، هپاتیت حاد و مزمن و سرطان کبد در هر دو روش خوشه بندی تقریباً یکسان گردید ولی درصد افتراق بیماری سیروز کبدی در روش خوشه بندی فازی به مراتب بیشتر از روش خوشه بندی کلاسیک شد، لذا برآورد احتمالی بیماران سیروز کبدی در روش خوشه بندی فازی دقیقتر از روش کلاسیک می باشد. همچنین در این روش احتمال تعلق بیمار به کلیه خوشه ها قابل احتساب است در حالی که در روش خوشه بندی کلاسیک این موضوع قابل احتساب نمی باشد. لازم بذکر است در میان روش های خوشه بندی کلاسیک، تنها روش k-میانه بیماریها را به درستی تشخیص داده و در موارد دیگر برخی بیماریها به درستی تشخیص داده شده و یا هیچ بیماری تشخیص داده نشد. همچنین شاخص ضریب افزاز دان یکسان بودن تقریبی نتایج خوشه بندی فازی و کلاسیک در افتراق بیماریها را تایید نموده و نمودار سایه نما نیز نشان داد که خوشه بندی صورت گرفته در هر دو روش خوشه بندی معقول می باشد، گرچه درصد مشاهدات بیشتری در خوشه بندی فازی بیشتر به چشم می خورد.

**بحث و نتیجه گیری:** در این مطالعه روش خوشه بندی فازی و کلاسیک جهت تشخیص اولیه بیماری های کبدی مورد مقایسه قرار گرفتند در صورتی که در مطالعات مشابه این مقایسه انجام نشده است. درصد افتراق بیماری سیروز کبدی در خوشه بندی فازی بیش از خوشه بندی کلاسیک گردید لذا استفاده از این روش جهت تشخیص اولیه بیماری های کبدی پیشنهاد می شود و برآورد احتمالی بیماران سیروز کبدی در روش خوشه بندی فازی دقیقتر از روش کلاسیک خواهد بود.

**واژه های کلیدی:** تحلیل خوشه ای فازی، بیماری های کبدی، پارامترهای آزمایشگاهی، FCM

<sup>۱</sup> Fuzzy C-means

## Abstract

**Background and Objectives:** Liver is the biggest gland and the most complicated organ which is responsible for more than 500 duties. Specialists have to do multiple experimental tests and examine a lot of data in order to diagnose liver diseases. This study is conducted using the algorithm fuzzy cluster to help specialists to diagnose liver diseases such as liver cancer, chronic hepatitis, acute hepatitis, fatty liver and liver cirrhosis. Finally, results are compared to classic clustering.

**Methods:** In this study, 1123 patients suffering from liver disease are examined and classified by the classic and fuzzy cluster analysis. Checklists are used to record the patient's variables.

**Results:** After classic and fuzzy cluster analysis has been done and percentage of differentiation of patients suffering from liver diseases has been compared using two clustering methods, it has shown that although differentiation of four liver diseases (fatty liver, chronic hepatitis, acute hepatitis and liver cancer) is similar in two clustering methods and leads to the same differentiation in diseases, percentage of cirrhosis is greater in fuzzy clustering than in classic clustering. Therefore, rough estimation of patients suffering from liver cirrhosis is more precise in fuzzy clustering than classic clustering. Also, in this method, it is possible to find out the probability to which group patient belongs but it is impossible in classic clustering. Additionally, partition coefficient index confirms the sameness of results of classic and fuzzy clustering methods in disease differentiation. Silhouette curve shows that clustering done in both methods is reasonable; although, observations are more compresses in fuzzy clustering.

**conclusion:** In this study, fuzzy clustering is compared to classic clustering in order to initially diagnose liver diseases. In previous studies this kind of comparison was not done. Although the percentage of differentiation of liver diseases including fatty liver, acute hepatitis, chronic hepatitis and liver cancer is almost similar in both clustering method, fuzzy clustering acts more precisely in cirrhosis because it shows greater differentiation than classic clustering. Therefore, usage of indices obtained from fuzzy and classic cluste analysis are advisable in diagnosis of liver diseases.

**Keywords:** Fuzzy cluster analysis, Liver diseases, Laboratory indices, FCM