



فرایند کاوی تخصیص و پرداخت وام

حسین علی بنده لو (۴۰۱۶۶۳۳۱۰۰۴)

استاد درس:

سرکار خانم دکتر الهام آخوند زاده

تابستان ۱۴۰۳

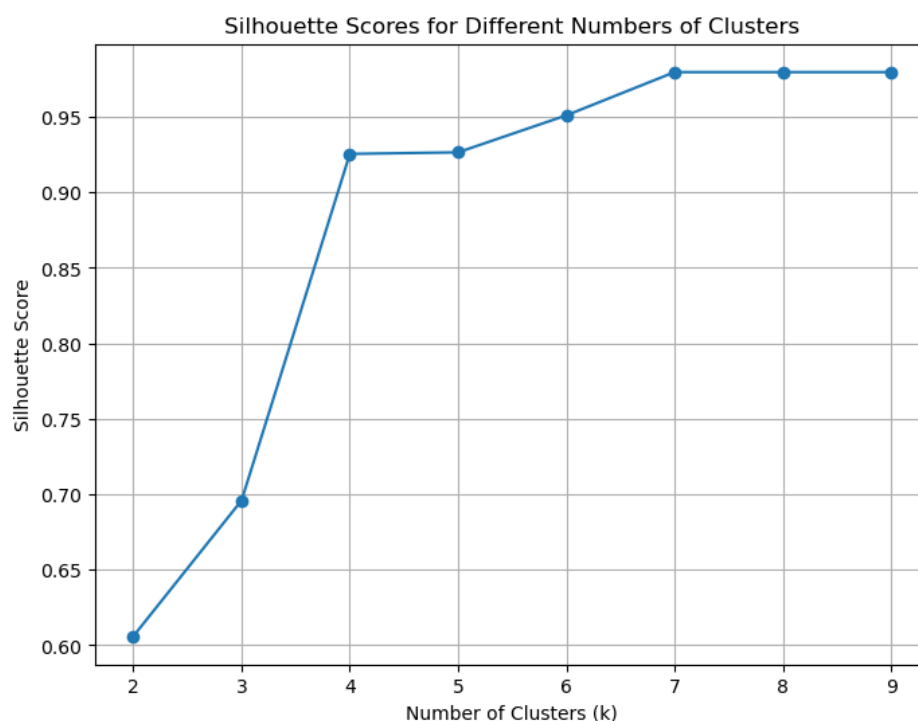
صفحه	عنوان
۲.....	مقدمه
۲.....	۱- الگوریتم K-Means
۲.....	۲- الگوریتم K-Modes
۳.....	۳- الگوریتم خوشه بندی سلسله مراتبی تجمیعی
۴.....	۴- تفسیر خوشه ها
۵.....	۴-۱- خوشه ۵
۵.....	۴-۲- خوشه ۳
۶.....	۴-۳- خوشه ۷
۶.....	۴-۴- خوشه ۶
۷.....	۴-۵- خوشه ۱
۷.....	۴-۶- خوشه ۴
۸.....	۴-۷- خوشه ۲

مقدمه

به منظور ارزیابی فرایندهای حال موجود و خوشه بندی آن ها، از سه الگوریتم K-Modes، K-Means و خوشه بندی سلسله مراتبی استفاده نمودیم. در هر روش از شاخص ارزیابی Silhouette به منظور ارزیابی خوشه بندی استفاده نموده ایم. همچنین در خوشه بندی سلسله مراتبی، از شاخص Cophenetic Correlation Coefficient برای انتخاب بهترین روش برای ادغام خوشه ها استفاده نموده ایم. نتایج نشان می دهد که هر سه روش بهترین تعداد خوشه ها را ۷ خوشه و با مقدار سیلوئت ۰.۹۸ بدست داده اند.

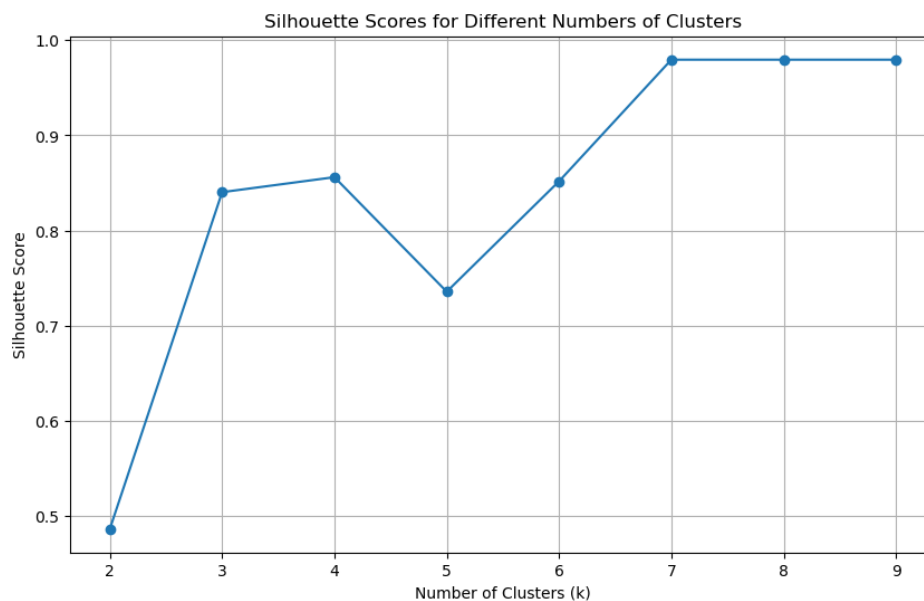
۱- الگوریتم K-Means

در این مجموعه داده با ۴۹ مورد تخصیص وام مواجه هستیم. ابتدا هر trace برای هر مورد از این ۴۹ مورد را بصورت دنباله ای از فعالیت ها در آورده سپس ماتریس فاصله با روش لونشتاین را محاسبه نمودیم. بر این اساس الگوریتم K-Means را بر روی این ماتریس فاصله اعمال کرده و با شاخص سیلوئت نتیجه هر خوشه را ارزیابی نمودیم. در شکل زیر، بهترین تعداد خوشه ۷ است.



۲- الگوریتم K-Modes

در این الگوریتم نیز بهترین تعداد خوشه ۷ بدست آمده و میانگین سیلوئت کل، برابر با ۰.۹۸ می باشد.

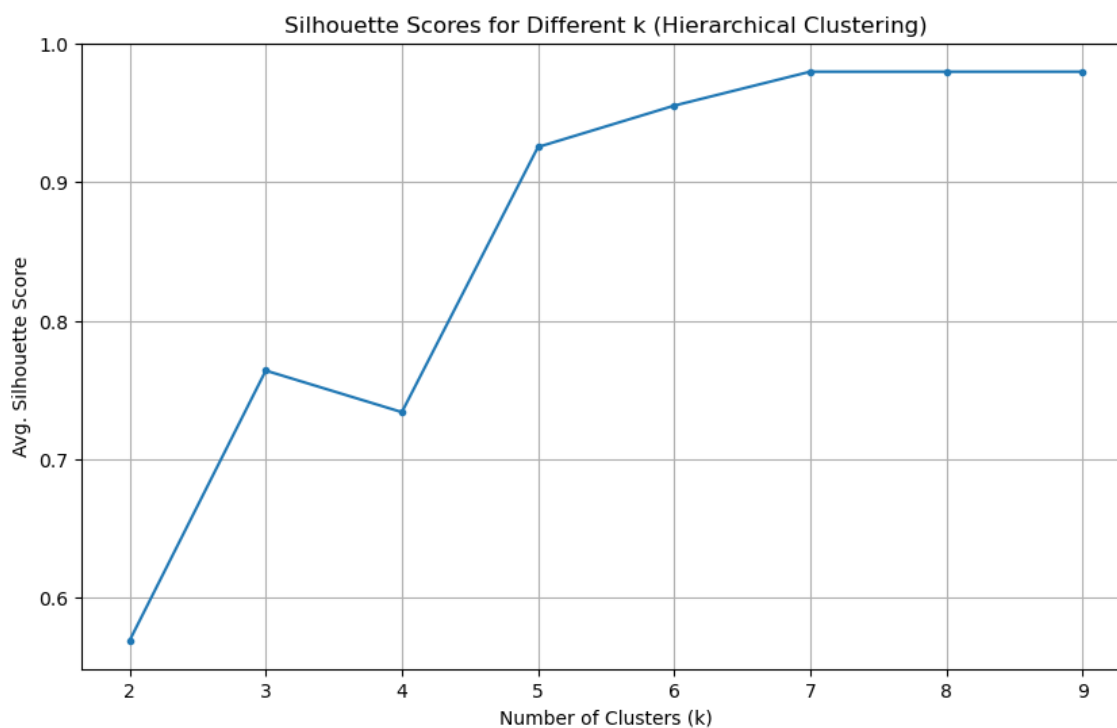


۳- الگوریتم خوشه بندی سلسله مراتبی تجمیعی

در ابتدا با روش‌های مختلف اقدام به انجام خوشه بندی سلسله مراتبی کرده و مشاهده می شود که روش میانگین، برای تجمیع خوشه های نزدیک به هم، بیشترین میزان شاخص Cophenetic Correlation Coefficient را به دست می دهد.

ردیف	روش	شاخص CPCC
1	single	0.847
2	complete	0.944
3	average	0.956
4	centroid	0.953
5	median	0.953
6	weighted	0.947
7	ward	0.833

حال با روش میانگین اقدام به خوشه بندی نموده و بهترین تعداد خوشه ها را با سیلوئت بدست آوردیم که مانند قبل، ۷ است.



۴- تفسیر خوشه ها

میانگین تعداد فعالیت ها	میانگین زمان تکمیل فرایند	درصد از کل (تجمیعی)	تعداد مورد	برچسب خوشه
2	0 days 00:00:44.62	45%	22	5
6	0 days 00:00:33.98	67%	11	3
6	1 days 14:54:55.93	86%	9	7
8	167 days 01:30:20.99	90%	2	6
12	3 days 23:39:35.06	94%	2	1
2	0 days 00:02:26.96	98%	2	4
16	0 days 02:02:45.60	100%	1	2

با توجه به جدول بالا، تعداد مورد های تخصیص وام، حدود ۸۶ درصد از کل موردها معادل ۴۲ مورد، در خوشه های ۵، ۳ و ۷ قرار گرفته اند. و سایر خوشه ها تعداد بسیار کمی را به خود اختصاص داده اند. مانند خوشه های ۱، ۴ و ۲.

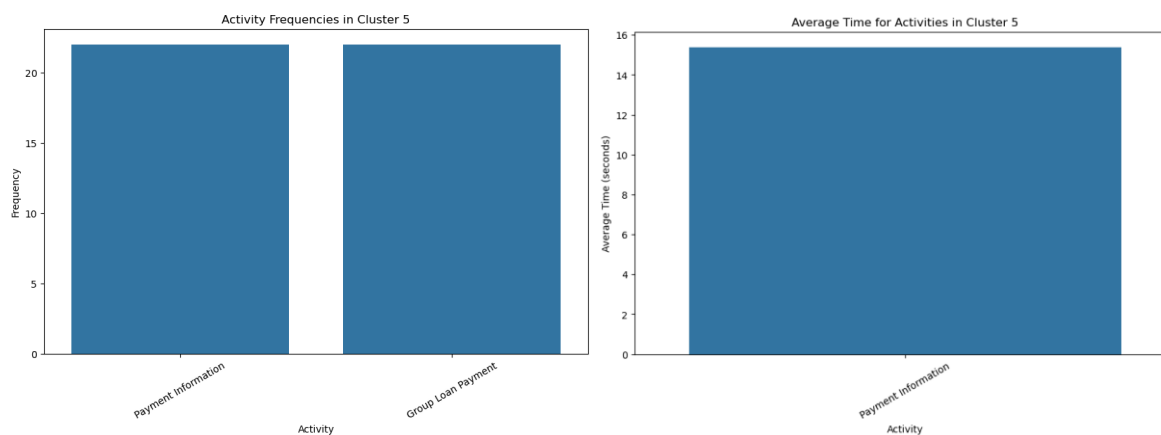
مشاهده می شود که خوشه های ۵ و ۳، میانگین زمان بسیار کمی را برای انجام فرایند داشته اند در حالی که خوشه ۷ به مدت ۱ روز و ۱۴ ساعت بطور میانگین زمان برده است. خوشه ۶ با اختلاف بسیاری از سایرین، با ۲ مورد از کل ۴۹ مورد تخصیص وام، ۱۶۷ روز طول کشیده تا انجام شود.

خوشه ۵ و ۴ با ۲ فعالیت به طور میانگین انجام شده اند. در حالی که خوشه های ۳ و ۷ با ۶ فعالیت و خوشه ۶ با ۱۲ فعالیت انجام گرفته اند. خوشه ۲ و ۱ به ترتیب با ۱۶ و ۱۲ فعالیت، خوشه هایی با بیشترین تعداد فعالیت هستند.

حال به ترتیب خوشه های جدول، اقدام به تحلیل فعالیت های هر کدام از خوشه ها می کنیم.

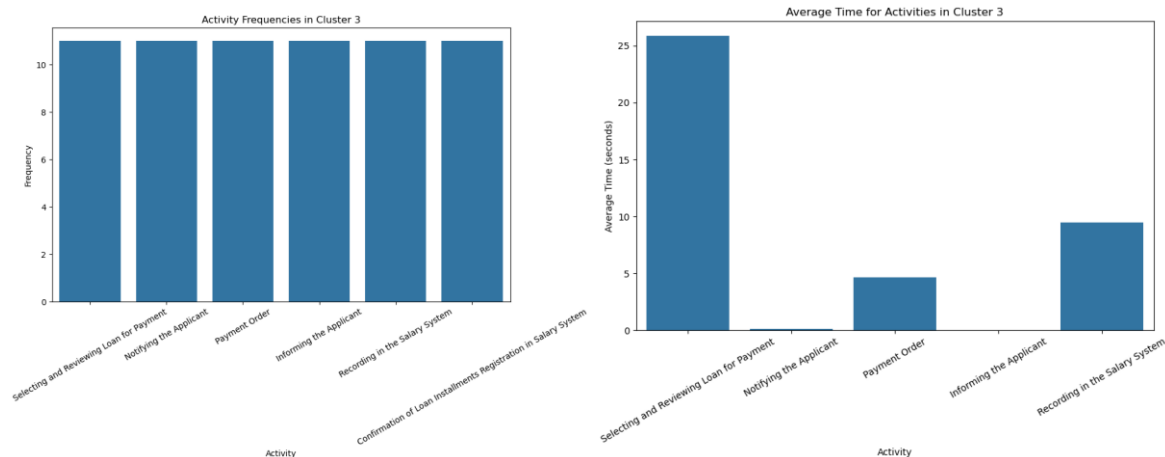
۴-۱- خوشه ۵

در این خوشه دو فعالیت دیده می شود که فعالیت اطلاعات پرداخت به طور میانگین ۱۵ ثانیه طول کشیده و سپس به طور گروهی به ۲۲ مورد پرداخت وام صورت گرفته است که موردی عجیب است و نیازمند بررسی دقیق است زیرا هیچ کدام از گام های قبلی برای درخواست متقاضی و بررسی پرونده و ... انجام نگرفته است و احتمالاً کارشناس مربوطه تقلب نموده است.



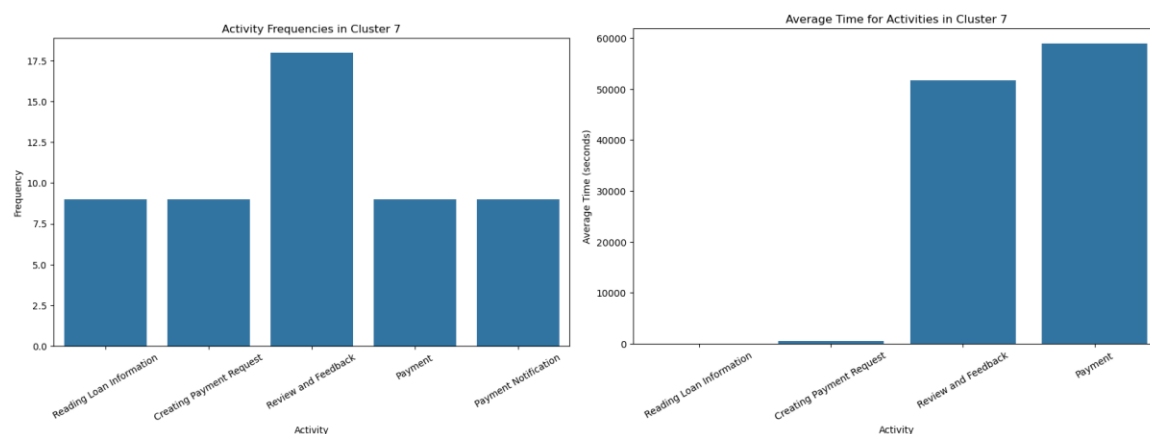
۴-۲- خوشه ۳

در خوشه ۳، فعالیت انجام گرفته است که هر فعالیت به تعداد یکسان میان ۱۱ مورد تخصیص وام صورت پذیرفته است. اما فعالیت اول تحت عنوان انتخاب و بررسی جهت پرداخت وام بیش از سایر فعالیت ها زمان برده است.



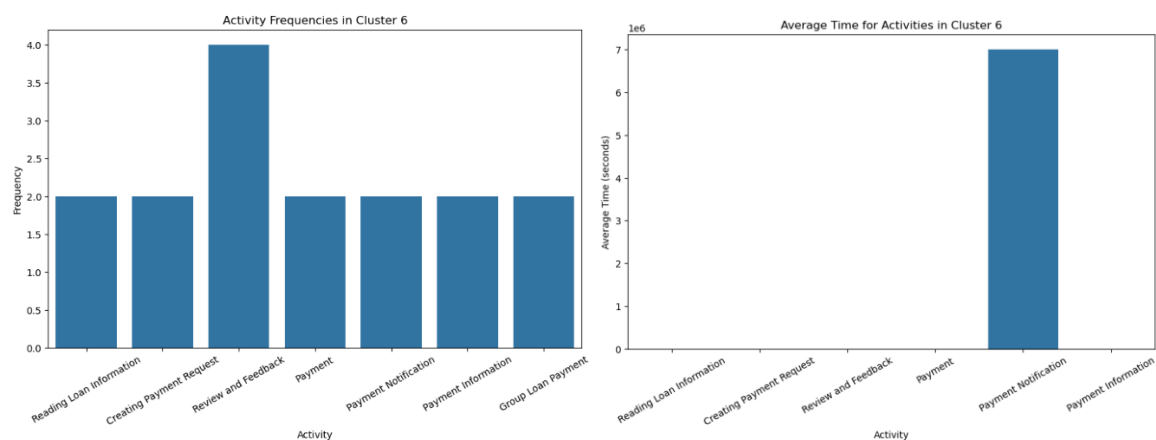
۴-۳- خوشه ۷

در این خوشه اما فعالیت بررسی و اعلام نظر تقریباً به تعداد دوبرابر سایر فعالیت ها تکرار شده و به نظر به علت نقص پرونده و مدارک مربوطه کارشناس به متقاضی بازخورد داده است. در این فرایند زمان بسیاری در فعالیت بررسی و اعلام نظر گرفته شده است و نیز پرداخت با تاخیر بسیاری همراه شده است.



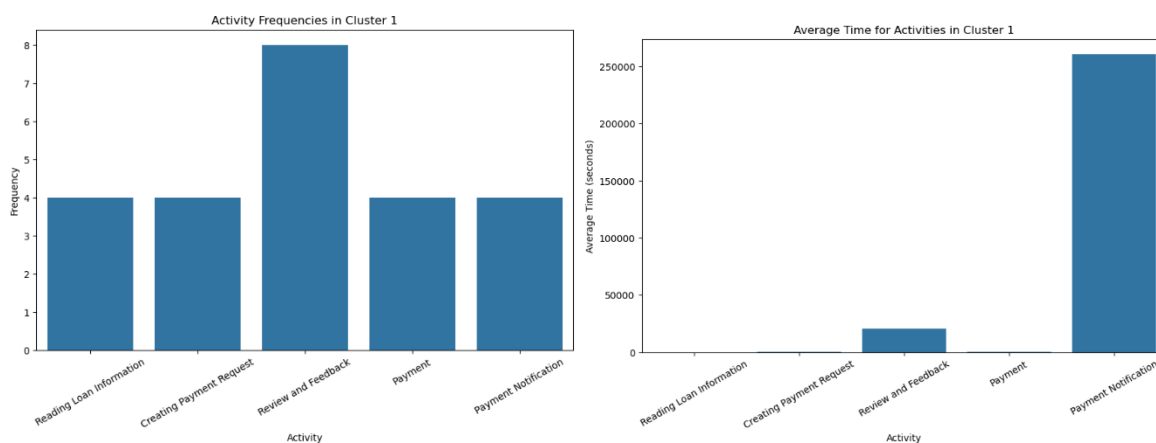
۴-۴- خوشه ۶

در این خوشه میانگین زمان انجام کل فرایند ۱۶۷ روز به طول انجامیده است و فقط ۲ مورد از کل ۴۹ مورد تخصیص وام را تشکیل می دهد. در این خوشه، تمامی فعالیت ها به غیر از اطلاع رسانی پرداخت با اختلاف چند ثانیه پس از دیگری انجام شده اما اطلاع رسانی پرداخت پس از مدت زمان زیادی انجام گرفته است.



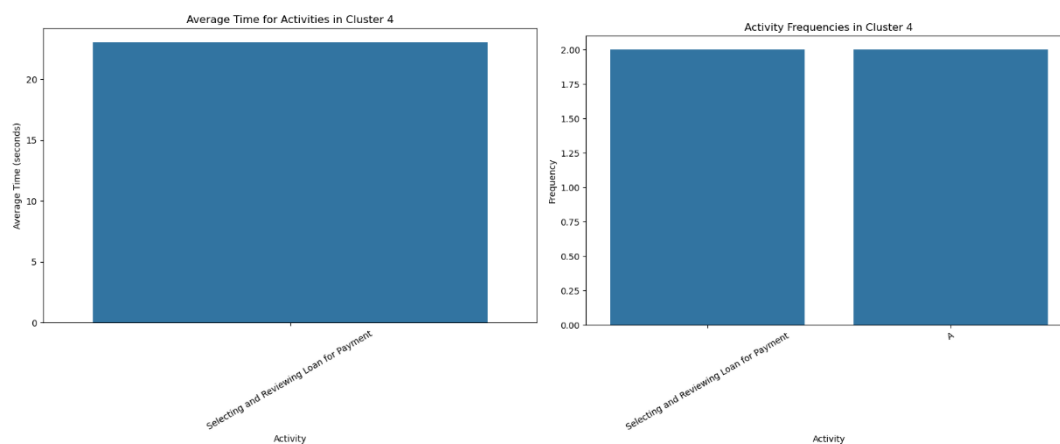
۵-۴- خوشه ۱

در این خوشه بیشترین زمان در فعالیت اطلاع پرداخت انجام شده و به طور میانگین ۱۲ فعالیت در این خوشه انجام گرفته است. هر فرایند برای هر دو موردی که در این خوشه قرار گرفته اند به طور میانگین ۳ روز طول کشیده تا تکمیل شود. همچنین بررسی و اعلام نظر، فعالیتی با بیشترین میانگین رخداد است.



۶-۴- خوشه ۴

این خوشه نیز همانند دو خوشه دو فعالیت دارد. فعالیت A قابل تفسیر نیست و فقط دو مورد در خوشه ۴ قرار دارند که در یک فرایند پرداخت وام به آن‌ها انجام گرفته است. این خوشه نیز مانند خوشه ۵ نیازمند بررسی دقیق تر است. این خوشه را می توان به عنوان داده پرت در نظر گرفت.



۴-۷- خوشه ۲

این خوشه که فقط یک مورد تخصیص و ام است، یک داده پرت است که ۱۶ بار فعالیت مختلف برای آن انجام شده و بیشتر فعالیت ها پیش از بررسی و اعلام نظر رخ داده اند. نکته عجیب آن است که کل این ۱۶ فعالیت در حدود ۲ ساعت انجام گرفته است و نیازمند بررسی است.

