

معرفی برنامه

برنامه این فاز در ادامه فاز قبل است. در این برنامه الگوریتم K Means پیاده شده است. در این پیاده سازی ابتدا K کلاستر با مراکز تصادفی ایجاد می شود. سپس به صورت تکراری نقاط را به کلاستری که نزدیک تر است انتساب می کنیم. پس از انتساب نقاط، مراکز هر کلاستر را به مرکز نقاط درون آن کلاستر جابه جا می کنیم. این مراحل به صورت تکراری ادامه پیدا می کند تا اختلاف مقدار SSE در دو مرحله اجرا کمتر از آستانه تعریف شده توسط کاربر شود و یا اینکه به حداکثر تعداد تکرار (که توسط کاربر مشخص شده باشد) برسیم. در هر تکرار مقدار SSE چاپ می شود. شکل زیر اجرای برنامه را نشان می دهد.

```
./main ./iris_bezdek.txt 3 100 0.001 2
```

```
Run #0
-----
Iteration #1: SSE = 170.442
Iteration #2: SSE = 94.1946
Iteration #3: SSE = 86.4247
Iteration #4: SSE = 84.4919
Iteration #5: SSE = 83.5791
Iteration #6: SSE = 82.727
Iteration #7: SSE = 81.5436
Iteration #8: SSE = 80.8064
Iteration #9: SSE = 79.8736
Iteration #10: SSE = 79.3444
Iteration #11: SSE = 78.9213
Iteration #12: SSE = 78.8557

Run #1
-----
Iteration #1: SSE = 347.337
Iteration #2: SSE = 193.897
Iteration #3: SSE = 114.693
Iteration #4: SSE = 92.6966
Iteration #5: SSE = 83.021
Iteration #6: SSE = 79.6157
Iteration #7: SSE = 79.012
Iteration #8: SSE = 78.8514
Iteration #9: SSE = 78.8514
```

پیاده‌سازی برنامه

نقاطی که از فایل خوانده می‌شوند در یک آرایه دوبعدی به نام `pointers` و با ابعاد $N \times D$ قرار دارند. نقاطی که برای مراکز کلاسترها انتخاب شده است نیز در یک آرایه دوبعدی با ابعاد $K \times D$ به نام `centers` قرار دارد. کلاستر هر نقطه هم توسط مقداری در یک آرایه یک بعدی به نام `cluster` مشخص می‌شود. یعنی مقدار `I` ام در این آرایه شماره کلاستر برای نقطه `I` ام خواهد بود.

در ابتدا مراکز به صورت تصادفی انتخاب شده است. سپس با توجه به این مراکز نقاط به کلاستر با نزدیک‌ترین مرکز منتسب می‌شوند. حال که نقاط به کلاسترها منتسب شده‌اند، مراکز کلاسترها به میانگین نقاط درون آن‌ها جابه‌جا می‌شود. سپس این مراحل از اول تکرار می‌شود.

در هر تکرار مقدار `SSE` محاسبه می‌شود. اگر تفاوت دو مقدار متوالی `SSE` کمتر از حد آستانه باشد. الگوریتم به تعادل رسیده است و پایان می‌یابد.

کل این مراحل به تعداد دفعاتی که توسط کاربر مشخص شده است تکرار می‌شود.