# پروژه ارشد هوش مصنوعی

عنوان: خوشه­بندی داده­ها روی نقشه

دانشجو: صدرا صمدی طاهرگورابی

شماره دانشحویی: 990197451

استاد: مجید ایرانپور مبارکه

دانشگاه: پیام نور واحد نجف­آباد

پاییز 1401

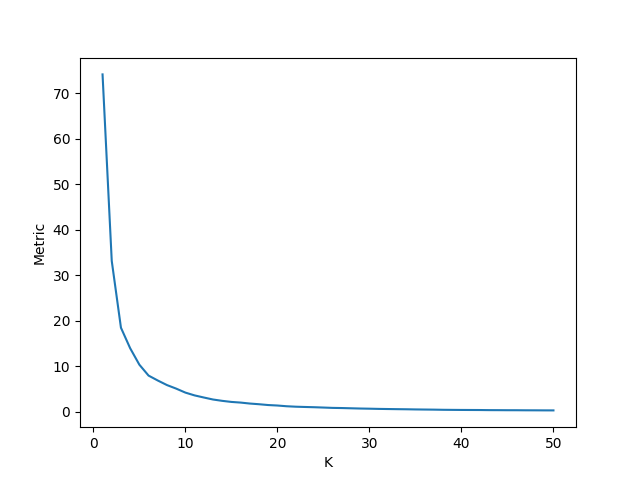
# توضیحات

در فاز اول این پروژه ما مشخصات مکانی مجموعه­ای از نقطه­ها در اختیار داریم که میخواهیم ابتدا آن­ها را در پنچره­های یک کیلومتر در یک کیلومتر جداسازی کرده و سپس روی این پنچره ها الگوریتم خوشه­بندی را پیاده­سازی کنیم.

این پروژه شامل دو بخش Backend و Frontend است.

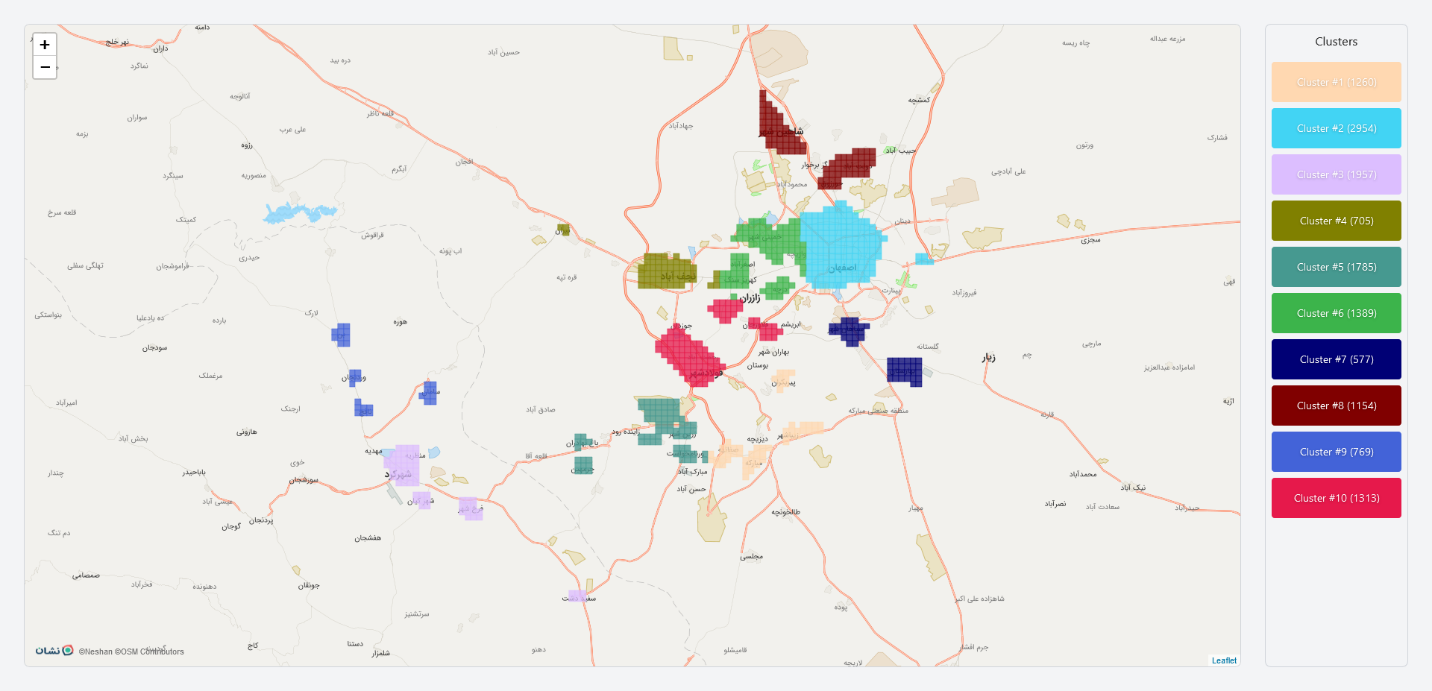
در بخش Backend که توسط زبان برنامه نویسی پایتون نوشته شده است مراحل اصلی یعنی پردازش داده­ها و اجرای الگوریتم خوشه­بندی انجام می­شود که به شرح زیر است:

1. ابتدا داده­های موجود در فایل ورودی را به صورت یک جدول (DataFrame) بارگذاری می­کنیم.
2. داده­های ناقص را حذف کرده و مشخصات مکانی دقبق هرکدام را به صورت Longitude و Latitude در ستون های مختلف قرار می­دهیم.
3. ناحیه­ای که نقطه­ها در آن قرار دارند را به پنجره­های یک در یک کیلومتر تقسیم کرده و مشخصات مکانی و تعداد نقاط موجود در هرکدام را مشخص می­کنیم.
4. از آنجا که نقطه­ها به صورت Longitude و Latitude هستند و این دو پارامتر مقیاس های متفاوتی دارند بهتر است که آن­ها را Normalize کرده و در بازه­ی صفر تا یک قرار دهیم.
5. الگوریتم خوشه بندی K-Means را با تعیین کردن حداکثر تعداد خوشه (K) روی پارامترهای مکانی نرمال شده پنجره­ها اجرا میکنیم تا مشخص شود هر خانه به کدام خوشه تعلق دارد.
6. برای یافتن تعداد مناسب خوشه ها از روش Elbow بهره می­بریم، یعنی به ازای تعداد خوشه­های مختلف الگوریتم را اجرا کرده، مجموع مربعات خطا (SSE) را محاسبه و در نمودار نمایش میدهیم. نقطه ای که در آن نمودار به شکل آرنج یک دست درآید معمولا بهترین انتخاب است:



1. نتیجه نهایی را در یک فایل خروجی ذخیره می­کنیم.

در بحش Frontend که به صورت یک وب اپ و به کمک کتابخانه ReactJS طراحی شده است، اطلاعات موجود در فایل خروجی تولید شده در قسمت قبل را با یک رابط کاربری ساده نمایش می­دهیم.



همانطور که در نمونه تصویر بالا مشخص است، پنجره های موجود روی نقشه نمایش داده شده و خوشه­های مربوط به هر کدام با رنگ­های مختلف متمایز شده­اند. با قرار دادن نشانگر موس روی هرکدام، تعداد داده های موجود در آن نمایش داده می­شود. همچنین تعداد کل داده­ها در هر خوشه نیز در سمت راست قرار داده شده­اند.

# راه­اندازی بخش Backend

برای این بخش نیاز است که نسخه 3.8 یا بالاتر نرم­افزار Pyhton و pip روی سیستم شما نصب باشد.

<https://www.python.org/>

کدهای اصلی این قسمت در مسیر /src/app/main.py و فایل­های ورودی و خروجی در مسیر /data قرار داده شده­اند.

جهت اجرا مراحل زیر را به ترتیب انجام دهید:

1. یک خط فرمان (command line, cmd, terminal, powershell, …) را دقیقا در مسیر پوشه /backend اجرا کنید.
2. دستورات زیر را جهت ساخت یک محیط مجازی و نصب پکیج­های استفاده شده در پروژه اجرا کنید:
3. python -m venv .venv
4. .\.venv\Scripts\activate
5. pip install poetry
6. poetry install
7. منتظر بمانید تا پکیج ها دریافت و نصب شوند (اولین اجرا ممکن است کمی زمان بر باشد).
8. جهت نمایش نمودار مجموع مربعات خطا دستور زیر را وارد کنید (این قسمت اختیاری است):
9. poetry run metrics
10. جهت اجرای الگوریتم خوشه­بندی و ذخیره اطلاعات در یک فایل خروجی دستور زیر را وارد کنید:
11. poetry run export
12. به طور پیش­فرض الگوریتم برای 10 خوشه اجرا می­شود. برای تغییر این مقدار می­توانید در انتهای دستور قبل تعداد موردنظر را هم وارد کنید، مثال:
13. poetry run export 15
14. اگر همه مراحل بدون مشکل اجرا شوند، یک فایل خروجی در مسیر /data/output.xlsx ساخته می­شود.

# راه­اندازی بخش Frontend

برای این بخش نیاز است که یکی از نسخه­های (ترجیحا به­روز) نرم­افزار Node.JS و npm روی سیستم شما نصب باشد.

<https://nodejs.org/en/>

کدهای اصلی این قسمت در مسیر /src/app.tsx قرار داده شده­اند.

جهت اجرا مراحل زیر را به ترتیب انجام دهید:

1. یک خط فرمان (command line, cmd, terminal, powershell, …) را دقیقا در مسیر پوشه /frontend اجرا کنید.
2. دستور زیر را جهت نصب پکیج های استفاده شده در پروژه اجرا کنید:
3. npm install
4. منتظر بمانید تا پکیج ها دریافت و نصب شوند (اولین اجرا ممکن است کمی زمان بر باشد).
5. فایل خروجی که در بخش Backend ساخته شد را در مسیر /frontend/public/output.xlsx کپی کنید.
6. دستور زیر را اجرا کنید:
7. npm run start
8. پس از چند ثانیه وارد آدرس زیر در مرورگر خود شوید:
9. <http://localhost:3000/>
10. اگر همه مراحل بدون مشکل اجرا شوند، رابط کاربری برنامه در صفحه مرورگر نمایش داده می­شود.
11. برای بستن برنامه در خط فرمان دکمه های Ctrl و C کیبورد را دوبار به صورت همزمان فشار دهید.