

گزارش پروژه نهایی هوش محاسباتی

پویا ترابی

بهمن ۹۸

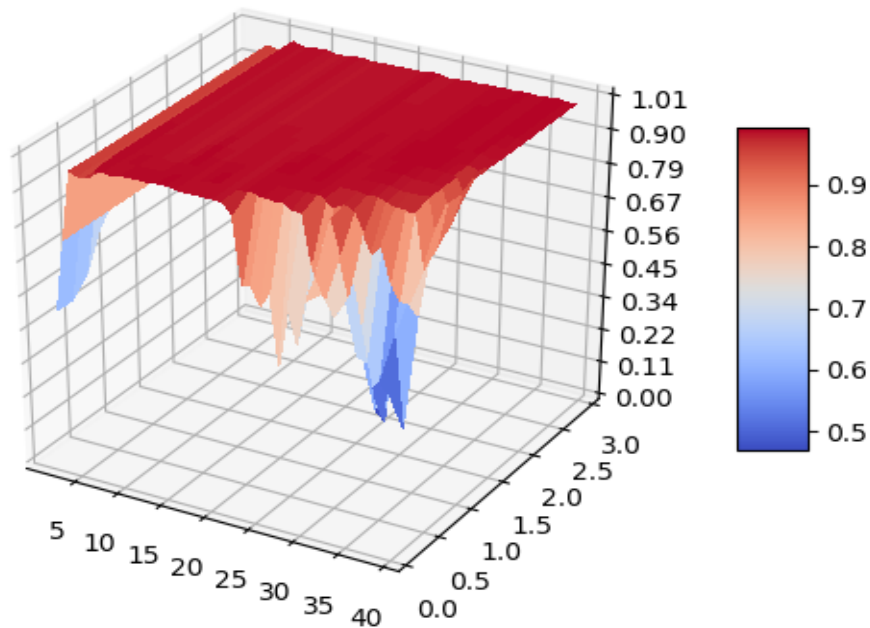
تعیین پارامترها

برای تعیین تعداد خوشه ها و اندازه شعاع بهینه، محدوده مقادیر زیر را امتحان کرده و دقت مدل را برای آن ها می سنجیم:

```
cluster_range = np.arange(2, 40, 1)
```

```
lambda_range = np.arange(0.1, 3, 0.2)
```

نمودار حرارتی دقت مدل برای داده های تست بر اساس پارامترهای بالا به صورت زیر است:



تحلیل:

برای تعداد خوشه بالاتر از ۱۰، دقت نمودار بسیار خوب است.

در بازه تعداد خوشه ۲۵ تا ۴۰، اندازه شعاع مرکز کمتر از ۱,۵ پارامتر نامناسبی بوده است. همان طور که در شکل پیداست در این محدوده ها، دقت تا مقدار ۷۰٪ افت داشته است.

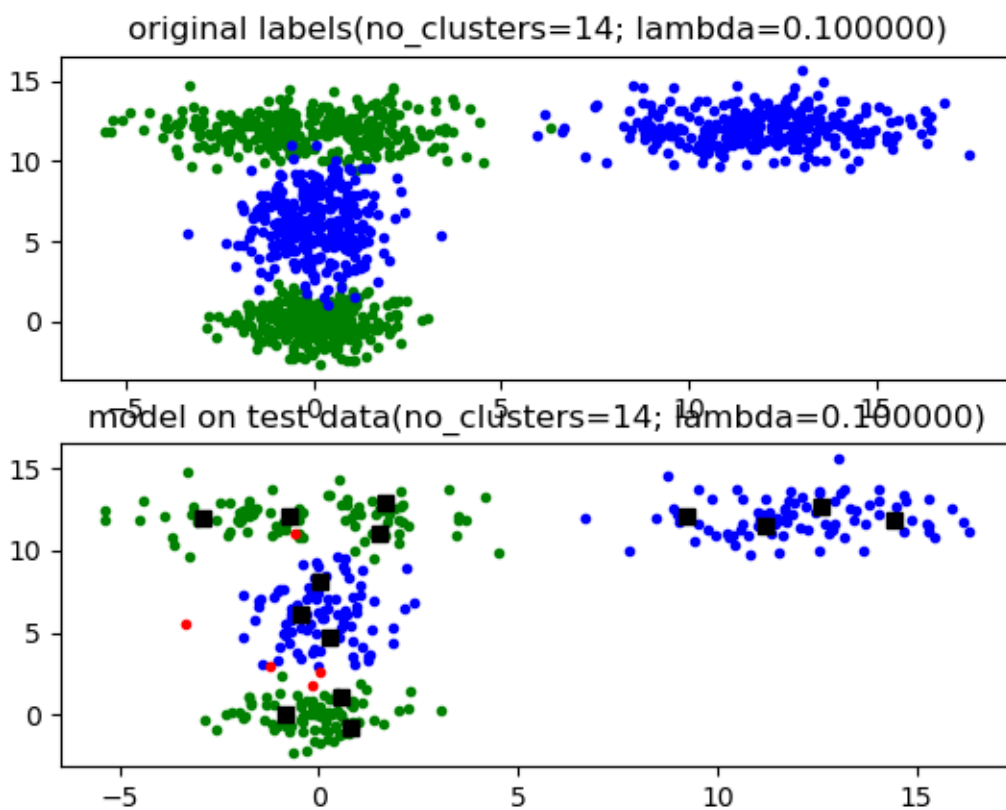
رسم نتایج

برای بهترین پارامترها، نمودار خروجی مدل رسم شده است و دقت مدل در زیر آورده شده است.

نقاط قرمز به معنی پیش بینی نادرست مدل است.

no_clusters=14; lambda=0.100000; accuracy on train data=0.980952

no_clusters=14; lambda=0.100000; accuracy on test data=0.986111



برای این خوشه بندی بهینه، مرزها مشخص شده اند. برای تعیین مرزها نقاط زیادی را به صورت توزیع یکنواخت تولید کرده و با توجه به مراکز بدست آمده، خوشه آن را مشخص می کنیم.

