|  |  |  |
| --- | --- | --- |
| C:\Users\rasoolamiri\AppData\Local\Microsoft\Windows\INetCache\Content.Word\th.jpg | **بسم الله الرحمن الرحیم** | ئئ |
| دانشگاه تهران |  | پردیس دانشکده­های فنی |

**یادگیری ماشین**

تهیه­ شده توسط:

**میلاد مسرور**

**فاطمه ساویی**

**سوال یک :**

بله مناسب میباشد و دقت کافی را دارند. همانطور که نمودارهای توان چند جمله ای 4 و 16 برای n-sample = 1000 نشان میدهند و همچنین score های به دست آمده برای این حالت ها، این نتیجه را میدهند که برای دادگان خطی استفاده از رگرسیون غیرخطی مناسب است.

زمانی که تعداد داده ها کم باشد. در این حالت مدل ما منجر به overfitting میشود.

**سوال دو :**

بیاییم بر روی دادگان خود، چند جمله ای هایی با توان های مختلف fit کنیم و برای هر مدل یک score گزارش بگیریم. هر کدام از مدل ها score بهتری نسبت به سایر مدل ها داشت، میتوان به این نتیجه رسید که توان چند جمله ای آن مدل بهترین توان برای دادگان ما میباشد. مثلا به نمودار scores vs n-samples موجود در کد دقت کنید، در این مورد NLR16 (چند جمله ای با توان 16) برای داده های ما مناسب میباشد.

**سوال سه :**

خیر مطلوب نمیباشد- بله امکان دارد – چون ممکن است مدل ما دچار اور فیتینگ شده باشد و زمانی که از این مدل برای دادگانی که تا به حال مدل ندیده است استفاده کنیم دچار خطا بسیار زیادی میشود.