**گزارش بخش چهارم تمرین دوم**

**محمد ویس مصطفی پور 97222085**

**سوال اول:** ابتدا یکی از کلاس ها را به عنوان کلاس اول انتخاب میکنیم، و بقیه ی کلاس های دیگر را به عنوان کلاس دوم در نظر میگیریم سپس روی آن لاجستیک رگرشن را اعمال میکنیم. سپس یک کلاس دیگر انتخاب میکنیم و همین مراحل را هم برای آن هم در نظر میگیریم. اگر ما n کلاس داشته باشیم در نهایت دارای N-1 مدل میشویم

**سوال دوم:** بله تفاوت احساس میشد و دقت مدل زمانی که آپ سمپلینگ انجام دادیم پایین آمد و زمانی که داده های ما نامتوازن بود نتایج بهتری به دست می آمد.

**سوال سوم:** بله تفاوت محسوس بود. دقت مدل بعد از forward selection افزایش پیدا کرده بود. چرا؟ چون ما در طی این فرایند فقط مهم ترین فیچر هایی که در تصمیم گیری تاثیر دارند را انتخاب کرده بودیم و فیچر هایی که ارتباطی به نتیجه گیری نداشتن حذف کردیم و این گونه دقت مدل بالا رفت

**سوال چهارم:**

**روش اول:** «بسته‌بندها» (Wrappers) شامل یک الگوریتم یادگیری به عنوان جعبه سیاه هستند و از کارایی پیش‌بینی آن برای ارزیابی مفید بودن زیرمجموعه‌ای از متغیرها استفاده می‌کنند. به عبارت دیگر، الگوریتم انتخاب ویژگی از روش یادگیری به عنوان یک زیرمجموعه با بار محاسباتی استفاده می‌کند که از فراخوانی الگوریتم برای ارزیابی هر زیرمجموعه از ویژگی‌ها نشات می‌گیرد. با این حال، این تعامل با دسته‌بند منجر به نتایج کارایی بهتری نسبت به فیلترها می‌شود.

**روش دوم:** «روش‌های توکار» (Embedded) انتخاب ویژگی را در فرآیند آموزش انجام می‌دهند و معمولا برای ماشین‌های یادگیری خاصی مورد استفاده قرار می‌گیرند. در این روش‌ها، جست‌وجو برای یک زیرمجموعه بهینه از ویژگی‌ها در مرحله ساخت دسته‌بند انجام می‌شود و می‌توان آن را به عنوان جست‌وجویی در فضای ترکیبی از زیر مجموعه‌ها و فرضیه‌ها دید. این روش‌ها قادر به ثبت وابستگی‌ها با هزینه‌های محاسباتی پایین‌تر نسبت به بسته‌بندها هستند. (منبع فرادرس)

**سوال پنجم:** عیب هردوی این روش ها این است که لزوما بهترین و تاثیرگذارترین فیچر را به ما تحویل نمیدهند.

راه حل چیست؟ راه حل استفاده ی همزمان از هردو روش است.

**سوال ششم:**

LDA چیست؟ تحلیل یا «آنالیز تشخیصی خطی» (Linear Discriminant Analysis – LDA) یک روش آماری برای کاهش ابعاد یک مسئله و تشخیص دسته‌ها بوسیله بیشینه‌سازی نسبت «پراکندگی بین گروه‌ها» (Scatters between groups) به «درون گروه‌ها» ( Scatters within groups) است.

در مسائل طبقه بندی روش LDA عملکرد بهتری از خود به جای میگذارد.

**سوال هشتم:**

MCC: پارامتر دیگری است که برای ارزیابی کارایی الگوریتم‌های یادگیری ماشین از آن استفاده می‌شود. این پارامتر بیان‌گر کیفیت کلاس‌بندی برای یک مجموعه باینری می‌باشد. ( MCC (Matthews correlation coefficient، سنجه‌ای است که بیان‌گر بستگی مابین مقادیر مشاهده شده از کلاس

باینری و مقادیر پیش‌بینی شده از آن می‌باشد. مقادیر مورد انتظار برای این کمیت در بازه 1− و 1 متغیر می‌باشد. مقدار 1+، نشان دهنده پیش‌بینی دقیق و بدون خطای الگوریتم یادگیر از کلاس باینری می‌باشد. مقدار 0، نشان دهنده پیش‌بینی تصادفی الگوریتم یادگیر از کلاس باینری می‌باشد. مقدار 1−، نشان دهنده عدم تطابق کامل مابین موارد پیش‌بینی شده از کلاس باینری و موارد مشاهده شده از آن می‌باشد. مقدار این پارامتر را به‌طور صریح، با توجه به مقادیر ماتریس آشفتگی به شرح زیر، می‌توان محاسبه نمود:

$$\mathrm{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$