**به نام خدا**

**پروژه درس شبکه عصبی**



**استاد: دکتر سادات ی**

**حل تمرین :**

**امیرمحمد غفار ی نسرین قربان ی**

**\*کمک گرفتن از دوستان و حل تمرین مجاز است ولی هرگونه کپی برداری از کد و نتایج دیگرا ن، به منزله نمره صفر خواهد بود.**

**پروژه اول)کلاس بندی با استفاده از شبکه های عصبی چند لایه(**

**پروژه تشخیص بیماران مبتلا به کبد چرب )MLP Binary classification/D(**

در این پروژه قسمتهای مختلف کد داده شده است و شما میبایست با ایجاد یک مدل شبکه عصبی چن دلایه به تشخیص افرادی که دارای این بیماری هستند بپردازی د.

* چالش اول: در مورد ویژگیهای داد ه شده یک دید کلی به دست آورید و یک پیشبینی از اینکه کدام ویژگ یها ممکن است تأثیر بیشتری روی خروجی داشته باشد را موردبررسی قرار دهی د.
* چالش دوم: یک پی شپردازش داده انجام دهید و نمون ههایی که دارای ویژگی با مقدار صفر هستند را موردبررسی قرار دهی د.)برای مثال فشارخون صفر در یک نمونه معنایی ندارد و باید با رو شهای معمول پی شپردازش اصلاح) Data cleaning( شود.
* چالش سوم: داد هها را با نسبت معقول به داد ههای تست و آموزش تقسیم کنید و سپس با استفاده از مدل standardscaler داد هها را استاندارد کنی د.
* چالش چهارم: مدل شبکه عصبی عمیق را با استفاده از ماژو لهای مورد علاقه خود

موردنیاز را با است دلال hyperparameter طراحی کنید و )Tensroflow/keras , pytorch( .خود تنظیم کنی د

* چالش پنجم: مدل خود را آموزش دهید و با رو شهای معمول مورد ارزیابی قرار دهی د )confusion

# .و خروجی کار را مصور کنی د )matrix , acc/loss plots

* چالش ششم: مدل آموزش داد هشده خود را درprediction مورد بررسی قرار دهی د.)یعنی یک ورودی تست را به این تابع ارجاع دهید و خروجی مدل را با خروجی واقعی مقایسه کنید(
* چالش هفتم: احتمالا دقت مدل به صد در صد نرسد و این یعنی ممکن است تعدادی از بیماران را به اشتباه سالم)false positive( و تعدادی از افراد سالم را به اشتباه بیمار نشان ده د)false negative(، به نظر شما کدام یک از این دو مورد، منجر به اتفاقات بد تری خواهد شد؟ برای حل کردن این مشکل چه پیشنهادی دارید؟
* چالش هشتم: برای یادگیری هر کتابخان های، بهتر است در وهله اول به اسناد آن مراجعه کنید. لطفا

# Early stopping ،callbacks api مراجعه کنید و راجع به دو<https://keras.io/api>به سایت .اطلاعات کسب کنید ModelCheckpoint

**پروژه دوم)رگرسیون با استفاده از شبکه های عصبی چندلایه(**

**پروژه پیشبینی قیمت خانه)MLP Regression/D (**

در این پروژه قسمتهای مختلف کد داده شده است و شما میبایست با ایجاد یک مدل شبکه عصبی چن دلای ه ب ه پیش بینی قیمت خانه ها با استفاده از ویژگی های داده شده بپردازی د. )توضیحات هر ویژگی در کد آورده شده است(.

چالش اول: ستون های id , unnamed را از دیتا ست حذف کنید و اگر علاوه بر این دو ویژگی ویژگ ی دیگری که به نظرتان در پیشبینی قیمت خانه موثر نیست را حذف کنید و سپس با استفاده از describe اطلاعات آماری دیتاست را بدست آورید.

چالش دوم: مهمترین بخش در کار های مربوط به شبکه عصبیData preprocessing میباشد در این دیتاست ممکن است برخی داده ها دارای ویژگی Null باشن د، آن ها را پیدا کنید و به جای داده های Null میانگین آن feature را قرار دهید.

چالش سوم: در مورد box plot تحقیق کنید و کاربرد آن را فرا گیرید و آن را با دو ویژگی price , waterfront مورد بررسی قرار دهید)برای راحتی میتوانید از کتابخانه seaborn استفاده کنید( چالش چهارم: با استفاده از تابع corr در کتابخانه pandas، وابستگی موجود در ویژگی ها را نسبت به price بدست آورید و نتیجه را با استدلال خود تحلیل کنید و نظر خود را راجع به این که شبکه از کدام ویژگی، برای یادگیری بیشتر استفاده میکند بیان کنید.

چالش پنجم: داد هها را با نسبت معقول به داد ههای تست و آموز ش و صحت سنج ی تقسیم کنید و سپس با استفاده از مدل standardscaler داد هها را استاندارد کنی د.

چالش ششم: مدل شبکه عصبی عمیق را با استفاده از ماژو لهای مور د علا قه خود

موردنیاز را با است دلال hyperparameter طراحی کنید و )Tensroflow/keras , pytorch( .خود تنظیم کنی د

چالش هفتم: مدل خود را تست و مورد ارزیابی قرار دهید و نتایج خود را مصور کنید)نمودار accuracy , loss شبکه حین آموز ش و برای داده های تست(

چالش هشتم: تنها ویژگی های sqft\_living , floors, waterfront را از دیتاست واکشی کنید و چالش های ششم و هفتم را تکرار کنی د، تفاوت های مشاهده شده را بیان کنید.

**پروژه سوم)کلاس بندی داده های تصویری( پروژه کلاس بندی لباس ها )CNN/ANN/D (**

در این پروژه قسمتهای مختلف کد داده شده است و شما میبایست با ایجاد یک مدل شبکه عصبی چن دلای ه کانوولوشنی و شبکه عصبی چند لایه)fully connected( مدلی را ایجاد کنید تا قادر به کلاس بندی لباس ها باشد .

همانطور که میدانید عکس ها در واقع آرای ههایی هستند که به نا به تعداد پیکسل ها، اعدادی از ۰ تا ۲۵۵ را شامل میشوند که در دیتاست داده شده ۷۰۰۰۰ عکس، از لباس ها آورده شده اند که لیبل مربوط به هر لباس در روی کد توضیح داده شده است.

لازم به ذکر است که ۶۰۰۰۰ از داده ها برای train , validation جدا شده اند و ۱۰۰۰۰ از آن ها برای test جدا شده اند. و مراحل preprocessing آن به خوبی انجام شده است

چالش اول: همه کتابخانه های خود را که فکر میکنید برای ساخت مدل کانوولوشنی خود مورد نیاز است فراخوانی کنی د

چالش دوم: سه مدل کانوولوشنی طراحی کنید که:

* مدل اول: حتما دارای ۴ لایه کانوولوشنی با Maxpooling باشد و دیگر هایپر پارامتر ها به صورت دلخواه انتخاب شون د، سپس شماتیک مدل نمایش داده شود و شبک ه train شود و مصور سازی های معمول ا نجام شده و دقت مدل روی داده های تست نمایش د اده شود )پیشنهاد:classification report در کتابخانه scikit-leran(
* مدل دوم: حتما دارای ۴ لایه کانوولوشنی با Meanpooling باشد و دیگر هایپر پارامت ر ها به صورت دلخواه انتخاب شون د، سپس شماتیک مدل نمایش داده شود و شبک ه train شود و مصور سازی های معمول انجام شده و دقت مدل روی داده های تست نمایش داده شود.
* مدل سوم: از بین دو مدل train شده، مدلی را که دارای دقت بیشتری روی داده های تست است انتخاب کرده و لایه dropout را نیز به آن اضافه کنی د، سپس شماتیک مد ل جدی د را نمایش دهی د و مدل را train کنی د، سپ س مصور سازی های معمول را ا نجام و دقت مدل روی داده های تست را نمایش دهید و تاثی ر لایه dropout را مورد بررسی قرار دهید.