

Alireza Hosseini      810100125      AI-HW4

بسمت اول : سوال اول :

(1) در صورت زیاد بودن داده های وجود نداشته حذف  
در غیر این صورت جایگزینی با الگوریتم های موجود مثل mean یا mode

(2) اختصاص وزن بیشتر به لایه های با داده کمتر  
تولید دیتا به روش های مثل GAN یا interpolation  
استفاده از روش های Ensemble

(3) normalize برای کاهش تأثیر نویز  
data Augmentation برای جلوگیری از overfitting به دلیل noise  
استفاده از الگوریتم های با قابلیت مدیریت بهتر نویز مثل SVM

(4) انتخاب یک feature بین چندین ویژگی که Correlation دارند و حذف بقیه  
کاهش بعد Feature Vector با PCA  
استفاده از مدل های robust نسبت به Correlation مثل Decision Tree

سوال دوم : برای امانت کردن تأثیر آزمون داین می تواند :

$$y = x \times \text{تعداد آزمون} + 60 + 5 \times \text{ساعت} = \text{نمره}$$

که به را می توان از طریق least squared بدست آورد. در روش least squared  
می شود ضرایبی که باعث حداقل شدن SSE می شود پیدا شود.  
در Gradient descent سعی می شود loss function حداقل شود. اگر SSE به عنوان  
loss function تعریف شود همان فارغی را می کند (به کمک مشتق ضمیمه)  
روش های دیگر : Normal Equation - Ordinary Least Square



prediction

span	300	30
ham	20	200
	span	ham

real

سوال سوم :

$$\text{accuracy} = \frac{TP + TN}{\text{total}} = \frac{500}{550} = \frac{10}{11} \approx 0.91$$

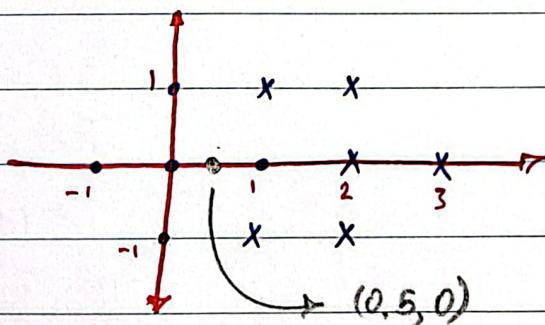
$$\text{precision} = \frac{TP}{TP + FP} = \frac{300}{330} \approx 0.91$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{300}{320} \approx 0.9375$$

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} = \frac{1.70625}{1.8475} \approx 0.923$$

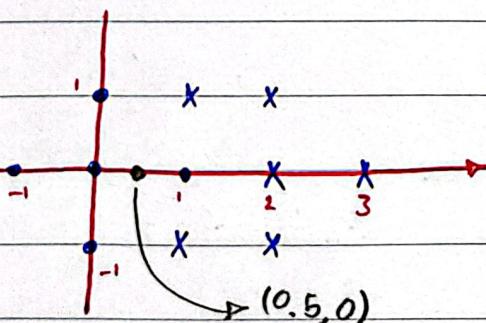
مسئله دوم : KNN

$K=3$ , distance metric = Euclidean



فاصله با دو نقطه 0.5  
برای نقطه سوم 4 کاندید وجود دارد  
ولی با هر الگوریتمی انتخاب کنیم  
باز در کلاس 2 می افتد

$K=3$ , distance metric = Manhattan

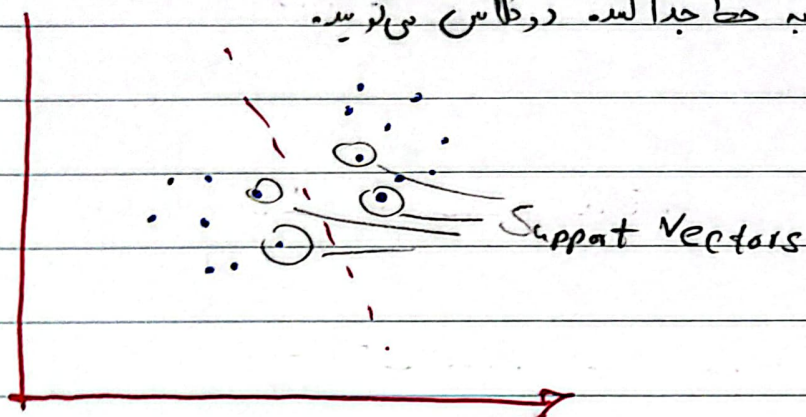


فاصله با دو نقطه 0.5  
برای نقطه سوم 6 کاندید وجود دارد  
ولی باز هم در هر صورت در کلاس 2 می افتد



## مبحث سوم : Support Vector Machine

• به نزدیکترین نقاط به خط جداکننده دو کلاس می‌گویند.



• SVM برای دیاست‌های خیلی بزرگ و با ابعاد (دیرنگی)‌های

زیاد مناسب نیست.

• کرنل‌ها نقاط را از فضای  $n$  به فضای دیگری می‌برند که در آن روابطشان تقریباً

خطی باشد، سپس این مسئله خطی با SVM حل می‌شود.

$$\phi: x \rightarrow \phi(x)$$

•  $hard\ margin$  نسبت به outlier ها بسیار حساس است و تمامی نمونه‌های هر کلاس

باید در یک طرف باشند، اما  $soft\ margin$  این مشکل را با اجازه دادن به وجود

تعدادی خطا حل می‌کند.



