HW3 Report

Section 1

Q1:

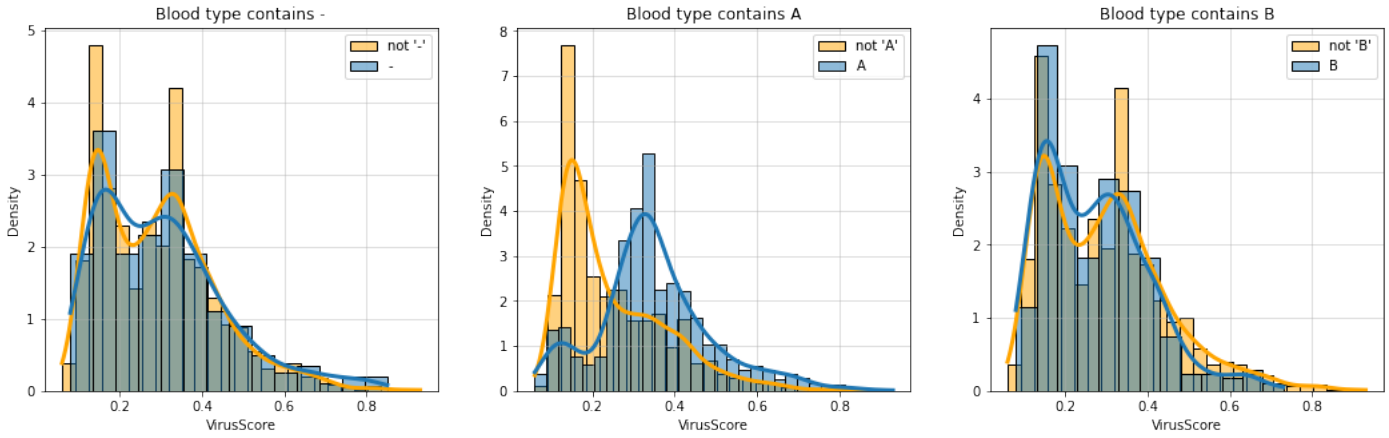

Figure 1: KDE plots of `VirusScore` conditioned different conditions of `blood_type`

Q2:

In figure 1 in the plot of A versus not A, we observe that the groups of patients with and and without "A" in their blood types are mostly seperable along a boundary that is approximately the `VirusScore` of 0.225.

Therefore, the condition of contains/does not contain A would be most informative for learning `VirusScore`. As it turns out, we decided already in hw1 to create this feature.

Q3:

$$\frac{\partial}{\partial b}\mathcal{L}\left(\underline{\mathbf{w}},b\right)\overset{a}{=}\frac{\partial}{\partial b}\frac{1}{m}\sum_{i=1}^{m}\left(\underline{w}^{\top}\underline{x}_i-b-y_i\right)^2\overset{b}{=}\frac{1}{m}\frac{\partial}{\partial b}\sum_{i=1}^{m}\left(\underline{w}^{\top}\underline{x}_i-b-y_i\right)^2\overset{c}{=}\frac{1}{m}\sum_{i=1}^{m}\frac{\partial}{\partial b}\left(\underline{w}^{\top}\underline{x}_i-b-y_i\right)^2\overset{d}{=}$$

$$\overset{d}{=}\frac{1}{m}\sum_{i=1}^{m}2\left(\underline{w}^{\top}\underline{x}_i-b-y_i\right)\cdot(-1)\overset{e}{=}\frac{2}{m}\sum_{i=1}^{m}\left(b+y_i-\underline{w}^{\top}\underline{x}_i\right)\overset{f}{=}\frac{2}{m}\left[mb+\sum_{i=1}^{m}\left(y_i-\underline{w}^{\top}\underline{x}_i\right)\right]\overset{g}{=}$$

$$\overset{g}{=}2b+\frac{2}{m}\sum_{i=1}^{m}\left(y_i-\underline{w}^{\top}\underline{x}_i\right)\Rightarrow\frac{\partial}{\partial b}\mathcal{L}\left(\underline{\mathbf{w}},b\right)=2b+\frac{2}{m}\sum_{i=1}^{m}\left(y_i-\underline{w}^{\top}\underline{x}_i\right)$$

a : Definition of $\mathcal{L}\left(\underline{\mathbf{w}},b\right)$
b : $\frac{1}{m}$ is scalar
c : Derivative of a sum is the sum of derivatives
d : Derivative of $\left(\underline{w}^{\top}\underline{x}_i-b-y_i\right)^2$ w.r.t b
e : 2 is scalar
f : sum of b
g : Removing b from the sum