גיל קוסטה
204022487

# APML EX 3

## PART 1

(1) נגדיר: $f_i(o_i;t)=\max\{t,o_i\}$

הוספת הפונ' למודל תגדיל להגד' את מרחב ההיפותזות
שהוא מגדיר כיוון שבדוגמה לפונ' ה- אוסף המקרים
הפונ' הנ"ל אינה לינארית לכן בהוספה למודל כלל שכבה
לינאריות הפונ' הנ"ל תרחיב את מרחב ההיפותזות לפונ'
שהן לא לינאריות, לכן התשובה לשאלה היא כן.

(2) $\sigma(x)=\frac{1}{1+e^{-x}}$

$$x \to \underset{f}{\boxed{e^{-\square}}} \to \underset{g}{\boxed{+}} \to \underset{h}{\boxed{\square^{-1}}} \to p, \qquad 1 \to g$$

נרצה לגזור $\frac{\partial p}{\partial x}$:

$$\frac{\partial p}{\partial x}=\underbrace{\frac{\partial p}{\partial h}}_{1}\cdot\underbrace{\frac{\partial h}{\partial g}}_{-\frac{1}{g^2}}\cdot\underbrace{\frac{\partial g}{\partial f}}_{1}\cdot\underbrace{\frac{\partial f}{\partial x}}_{-e^{-x}}=\frac{-1}{(1+e^{-x})^2}\cdot(-e^{-x})=\frac{e^{-x}}{(1+e^{-x})^2}$$

(3) נאתחל את משקולות ה-toy convnet באפסים במהלך ה-SGD

נרצה לחשב:

1) $\frac{\partial l}{\partial u}=2(p-y)\cdot m^T \rightarrow$ בעת שאנחנו מאתחלים את $u=\vec{0}$
נקבל כי $m=\vec{0}$ ולכן $\frac{\partial l}{\partial u}=0$

2) $\frac{\partial l}{\partial w_i}=\frac{\partial l}{\partial p}\cdot\underbrace{\frac{\partial p}{\partial m}}_{=u^T}\cdot\frac{\partial m}{\partial n}\cdot\frac{\partial n}{\partial o}\cdot\frac{\partial o}{\partial w_i} \rightarrow$ כיוון ש- $u=\vec{0}$ כל המכפלה
תתאפס ונקבל $\frac{\partial l}{\partial w_i}=\vec{0}$

לכן לא נשנה לעדכן את $w_1,w_2,u$ באיטרציה הראשונה, מכיוון
דבר לא ישתנה והתוצאה תהיה זהה גם באיטרציה השנייה
כלומר לא נצליח להתקדם באלגוריתם.