Abdelrahman Rezk

AOU University

NLP & ML student

&

Web Developer

بسم الله الرحمن الرحيم

الملخص ده بإذن الله لازم تكون شوفت قبليه تالت إسبوع من Andrew Ng

<https://www.coursera.org/learn/machine-learning>

وفيه مقابل ليه بالعربى مهندس هشام وده بصراحة ربنا يباركله ويجازيه خير على العلم الى بيقدمه والمساعده الجميله وده برده تالت اسبوع ممثل فى الجزء الرابع فى الكورس فى البلاى ليست بتاعته.

<https://www.youtube.com/watch?v=e1lAgjFsobc&list=PL6-3IRz2XF5Uq7PkI_PWOm_DLC2CFvSzU>

المذكور هنا هايكون جهد شخصى ومحاوله إنى اساعد فى تبسيط المصطلحات والمعادلات مع حل امثله عملية تربط جميع الاجزاء المختلفه خلال الإسبوع التالت بإذن الله فالمفروض إنك تكون شوفته.

"وهاكون شاكر لو كان هناك دعوة بظهر الغيب"

Classification

ببساطة التصنيف هو إنى بدل ما كنت بحاول اتوقع فى ال linear regression continuous value لا هنا انا بحاول اتوقع discrete value يعنى value محدده ممكن يكون في اكتر من value بحاول اتوقعها زى الارقام من 0 ل 9 فى الارقام ولكن فى الغالب بيكون معظم التصنيف binary بمعنى توقع الميل المرسل spam or not .

عشان ققدر أحقق حاجه زى كده وانى اخلى ال output بتاعى يكون قيم محدده بستخدم ال sigmoid function وهى بتعمل mapping للأرقام بين 0≤hθ​(x)≤1 بمعنى انى فى الاخر ال value الى هاتطلع هاتكون فى الرنج بين 0 و 1 وانا من هنا ببتدى احدد بقا لو كان مثلا اكبر من او يساوى .5 ققول انه spam اقل من كده ققول انه not-spam .

|  |  |
| --- | --- |
| output y | input x |
| 4 | **0** |
| 7 | **1** |
| 7 | **2** |
| 8 | **3** |

m = training examples

عدد الامثله الى هاقدر استخدمها عشان اعمل training للموديل بتاعى فى المثال الى فوق 4

n = number of features

الحاجات المميزه الى بتساعدنى اعمل predict كويس وهنا معنديش الا 1 بس الى هو x

y = actual value or output value

انا فى الاخر بتعامل مع supervised learning فبحاول انى اعمل mapping from input to output معين وده الى انا بخلى الموديل بتاعى يدرب عليه ويحاول يقلل الفرق قدر الامكان بين القيم المتوقعة والقيم الحقيقية.

y^ = h(x) = predicted value لكن هنا على عكس ال linear regression انا هنا هاستخدم معادله ال sigmoid .

ديه القيمه المتوقعة

∑ summation sign

Θ thetas parameters

قيم ال thetas الاخيره هى الى بتستخدم عشان احاول اعمل predict لحاجه جديدة مشفتهاش قبل كده عن طريق قيم ال thetas الى انا طلعتها من ال training examples .

α alpha called learning rate

ديه الى بتاثر فى تغير تقليل نسبه الخطأ بصورة كبيره او صغيره على حسب قيمتها.

**المعادلات المستخدمة**

Our new form uses the "Sigmoid Function," also called the "Logistic Function"

إحنا بنحاول إننا نشتغل بنفس معادله ال linear عد إننا بنيجى بقا مع ال h(x) ونستخدم ال sigmoid function والى هى بتخلينى اعمل mapping لل real number فى range(0,1) بس هنا ال h(x) هاتدينا ارقام بين ال (0,1) كانها probability وإحنا المفروض نعمل mapping بقا لو كان مثلا h(x) = .7 فده اقرب لل 1 وهكذا.

Cost function of logistic regression

J(*θ)* =

Vectorized version of cost function of logistic regression

Gradient Descent of logistic regression

نفس ال linear regression لكن ال h(x) هنا بتاعت ال sigmoid function

Vectorized version of Gradient Descent of logistic regression

**Another vectorize version**

**Binary Classification examples with 2 features**

**Dataset**

|  |  |  |
| --- | --- | --- |
| **Y0** | **X2** | **X1** |
| **0** | **78** | **34** |
| **0** | **43** | **30** |
| **0** | **72** | **35** |
| **1** | **86** | **60** |
| **1** | **75** | **79** |

**افترض انى theta 0=-25, theta1 =.20, theta1 = .20**

**والمفروض X بدل ما تبقا 5\*2 هاتبقا 5\*3 وهافترض vector X0 = 1 فالبتالى x هاتكون**

|  |  |  |
| --- | --- | --- |
| **X2** | **X1** | **X0** |
| **78** | **34** | **1** |
| **43** | **30** | **1** |
| **72** | **35** | **1** |
| **86** | **60** | **1** |
| **75** | **79** | **1** |

اولا cost function of vectorized version

M=5 training examples

الاول نجيب ال g الى هى ال sigmoid function والى كانت بتساوى

z = Xθ

بما إن X = 5\*3 و θ بتساوى 3\*1 فينفع يحصل بينهم multiplication والناتج هايكون vector 5\*1 .

z = =

h =

[ \* log() + \* log( 1- )] =

ضرب ال vector y = 1\*5 فى ال vector h = 5\*1 الناتج هايكون 1\*5 , 5\*1 هايكون 1\*1 الى هو -0.0179046

-0.0179046 + -0.098570968

هنا لما كانت ال y = 0 هاتلاحظ فى اول جزء من المعادله محسبناش الى اخر رقمين الى كانت فيهما ال y=1 والباقى كله طار عشان ضربنا فى صفر والجزء التانى العكس لما كانت ال y=0 بقا 1-0 بيساوى 1 والواحيد هى الى طارت.

J(θ)= 0.023295

ثانيا Gradient Descent of vectorized version

***ال sigmoid function الى إحنا كنا جايبنا فوق الى هى h***

***بما إن لما عملنا transpose لل X هاتبقا 3\*5 لما تضرب فى 5\*1 الى هو ناتج طرح ال فهيكون عندى vector 3 \* 1 يضرب فى 1/5 ويكون ده قيم ال theta الجديدة الى هى هاتسعدنى بعد كده اعمل predict عن طريق ال sigmoid function لما اخلص training بعد طبعا ما اكون حاولت انى ققلل ال cost function قدر الإمكان.***

=

يتم التكرار حتى تقليل قيمه ال error لل global minimum .

**بإذن الله الجاى هايكون التطبيق على Neuron Network واتمنى من الله اكون قدرت اساعد بحاجه بسيطة وإن شاء الله ققدر اكمل باقى الكورس على خير.**