



دانشگاه تهران - دانشکده مهندسی برق و کامپیوتر
سیستم های هوشمند - نیمسال اول 1398-1399
تمرین 1: بهینه سازی و طبقه بندی
مدرس: دکتر رشاد حسینی

1- تابع زیر را در نظر بگیرید.

$$f(x) = 1.5x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + 0.5x_1^4$$

تمام نقاط ایستای تابع را به دست آورده و نوع آن را مشخص کنید.

(10نمره)



2- مجموعه داده های یادگیری $S = \{(x_i, y_i) \in R^3 \times R \mid i = 1, 2, 3\}$ را در نظر بگیرید :

$$(x_1, y_1) = ((2, 0, 0), 1)$$

$$(x_2, y_2) = ((0, 1, 0), -1)$$

$$(x_3, y_3) = ((0, 0, 0.5), 1)$$

فرض کنید می خواهیم طبقه بند خطی پیدا کنیم: $f_w = \langle w, x \rangle$ با بردار وزن: $w \in R^3$

برای هر یک از حالت های زیر گرادینان را محاسبه کرده ، کاربرد هر یک از تابع های هزینه را بیان کنید . (در هر قسمت تابعی که باید آن را بهینه کنید نوشته شده است)

الف: Square loss with no Regularization

$$l(y^{\wedge}, y) = (y^{\wedge} - y)^2$$

ب: Square loss with l_1 Regularization

$$(y^{\wedge}, y) = (y^{\wedge} - y)^2$$

$$R(w) = ||w||_1$$



ج: Logistic loss with no Regularization:

$$l(y^{\wedge} و y) = \log(1 + e^{-y^{\wedge} \cdot y})$$

د: Logistic loss with l_2 Regularization:

$$l(y^{\wedge} و y) = \log(1 + e^{-y^{\wedge} \cdot y})$$

$$R(w) = \frac{1}{2} ||w||_2^2$$

(10نمره)



دانشگاه تهران - دانشکده مهندسی برق و کامپیوتر

سیستم های هوشمند - نیمسال اول 1398-1399

تمرین 1: بهینه سازی و طبقه بندی

مدرس: دکتر رشاد حسینی

3- تابع زیر را در نظر بگیرید :

$$f(x_1, x_2) = 3x_1^2 + 12x_1 + 8x_2^2 + 8x_2$$

از نقطه اولیه (1,1) شروع کنید. جهت گرادیان نزولی را پیدا کنید.

سپس با جست جوی خطی مقدار بهینه طول گام (α) را بیابید. (روش تحلیلی) با استفاده از جهت گرادیان و طول گام ، نقطه بعدی را بیابید .

سپس با استفاده از گرادیان کاهشی نقطه کمینه را محاسبه کرده و نتایج را مقایسه کنید.

برای هر الگوریتم دو بار تکرار محاسبه کنید.

(10نمره)



4- از کاربرد های مهم روش های بهینه سازی، استفاده آنها در کلاس بندی داده ها است. در این نوع مسائل، تعدادی داده داریم که دسته آنها مشخص است و هدف تعیین کلاس سایر داده ها است. هر داده توسط یک بردار ویژگی مشخص می شود و یک تابع هدف نیز بر اساس آن ها تعریف میشود. با بهینه سازی این تابع، معمولاً یک بردار وزن در مقابل بردار ویژگی داده ها بدست می آید. که توسط آن میتوان کلاس بندی داده های جدید را با تقریب خوبی مشخص کرد. در این سوال، تعدادی داده واقعی و تابع هدف مناسب برای این دادهها به شما داده می شود. شما با کمینه کردن تابع، مقدار بهینه پارامتر های تابع که همان وزن ها هستند را می یابید. تابع هدف زیر را در نظر بگیرید:

$$g(w) = \lambda ||w||^2 + \frac{1}{n} \sum_{i=0}^n \max(0, 1 - y_i(w \cdot x_i - b))$$

هدف از بهینه سازی یافتن پارامتر های w و b می باشد.

برای یادگیری از مجموعه داده های Fashion MNIST استفاده میکنیم که دارای 10 کلاس است. با استفاده از الگوریتم one vs. all این 10 کلاس را از یک دیگر جدا کنید.

مقدار مناسب پارامتر λ را از 10^{-10} تا 10^{10} با گام های 2 در توان (10^{-10} و 10^{-8} و 10^{-6} و ...) تغییر دهید و با 5-fold cross validation وزن ها را به دست آورید و نتایج را مقایسه کنید. (داده های یادگیری را به 5 قسمت تقسیم کرده، در هر بار یک مقدار برای λ قرار داده، با استفاده از 4 قسمت از داده ها، ضرایب را به دست آورده و روی قسمت 5 ام اعمال می کنیم)

نکات:

مقدار اولیه پارامتر w را یک عدد تصادفی از توزیع گاوسی با واریانس 0.1 در نظر بگیرید.



دانشگاه تهران - دانشکده مهندسی برق و کامپیوتر

سیستم های هوشمند - نیمسال اول 1399-1398

تمرین 1: بهینه سازی و طبقه بندی

مدرس: دکتر رشاد حسینی

الف- ابتدا گرادیان تابع را حساب کنید. سپس با روش نزول گرادیان تصادفی مقدار بهینه متغیرهای تابع را بیابید. شرط توقف الگوریتم را توضیح داده و مقدار تابع را به ازای توالی های X رسم کنید.

ب- مقدار طول گام را یکبار به صورت ثابت و بار دیگر به صورت نزولی انتخاب کنید و با تغییر پارامترهای مربوطه رفتار مشاهده شده را گزارش کنید.

ج- نمودار $loss$ را بر حسب تعداد تکرار برای هر بار رسم کنید.

د- به دلخواه برای یکی از حالت های $steepest descent$ را پیاده سازی کرده و نتیجه را با قسمت قبل مقایسه کنید.

(30نمره)



5- یکی از الگوریتم های کلاسیک برای شناسایی label الگوریتم KNN می باشد. این الگوریتم برای دو منظور استفاده می شود که یکی برای Classification و دیگری برای به دست آوردن توزیع (Distribution) می باشد. در درس سیستم های هوشمند (و این تمرین) با مورد اول کار خواهید کرد. داده های Fashion MNIST به شما داده شده است که شامل فایل های Train و Test به همراه Label های آن می باشد. هدف این سوال این است که شما با استفاده از الگوریتم غیر پارامتری KNN , Label داده های موجود در فایل Test را به دست بیارید.

الف) به ازای K های مختلف داده های داده شده را طبقه بندی کنید $k = \{2, 5, 10, 15\}$ (تعریف فاصله را Norm 2 فرض کنید)

ب) با توجه به قسمت قبل به نظرتان اگر مقدار K را خیلی بزرگ یا خیلی کوچک فرض کنیم چه نتیجه ای خواهد داشت؟ (دلایل خودتان را ذکر کنید)

ج) در الگوریتم KNN تعریف های متفاوتی از فاصله وجود دارد از جمله Euclidean distance و City Block distance و ... به ازای K بهینه بدست آمده در قسمت الف (آن K که بیشترین Accuracy را دارد) این بخش را با استفاده از فاصله های $d_1(x, y)$ و $d_2(x, y)$ تکرار کنید.

$$d_1(x, y) = \max |x_i - y_i| \quad d_2(x, y) = \sum_{i=1}^n |x_i - y_i|$$

(20نمره)



6- می خواهیم رمزی را که براساس حروف الفبای انگلیسی تنظیم شده است حدس بزنیم. با توجه به تعداد حروف (26 حرف کوچک ، (مانند a) و 26 حرف بزرگ (مانند A) و فاصله (در مجموع 53 کاراکتر برای هر موقعیت داریم. حال فرض می کنیم که رمز به صورت جمله باشد مانند

(The pen is mightier than the sword) که در مجموع دارای 34 کاراکتر است، تمام حالات ممکن برای حدس رمز صحیح برابر است با : $53 * 53 * 53 * \dots * 53$ (34 بار) که عدد 53^{34} بار حدس زدن برای رسیدن به پاسخ را تشکیل میدهد. واضح است که حدس چنین رمزی با تعداد حالات ممکن با روش های معمولی غیر ممکن است و می بایست از روش های فراابتکاری برای حل آن استفاده کرد.

حل مسئله حدس رمز توسط روش های فرا ابتکاری :

تابع هدف را برابر با تعداد فضاهای مشابه در رمز اصلی و رمز حدس زده شده در نظر گرفته و با یکی از روش های فراابتکاری مسئله را حل کنید. به طور مثال اگر رمز "hello" باشد و رمز حدس زده شده "Hello" تابع هدف برابر با 3 خواهد بود.

رمز "you can crack it" را با تعداد فضاهای 16 عدد در نظر بگیرید. برای هر فضا نیز 53 کاراکتر داریم. ابتدا اعداد تصادفی در بازه [0,52] تولید کرده تا برداری از فرض های اولیه داشته باشیم. سپس تابع هدف را محاسبه کرده و با کمک یکی از الگوریتم های ژنتیک یا تبرید شبیه سازی شده ، رمز صحیح را پیدا کنید.

(20نمره)



نکات:

- توجه کنید که نیمی از نمره تمرین مربوط به گزارش می باشد و لازم به ذکر است که رعایت اصول نگارشی حائز اهمیت می باشد.
- گزارش تمرین حتما باید به فرمت PDF باشد.
- نحوه نام گذاری تمرین بر اساس studentnumber_lastname_firstname.pdf باید باشد.
- زبان پیاده سازی تمرین می بایست Python یا Matlab باشد.
- هر گونه پرسش پیرامون تمرین را با ایمیل های aminfadaei116@gmail.com , negin.azizi96@hotmail.com , aliparchekan@gmail.com مکاتبه نمایید.