به نام خدا

سینا کریمی

9931050

تمرین دوم

سوال 1)

الف)

ابعاد ماتریس حاصل برابر با مقدار زیر خواهد بود:

حال ضرب ماتریس ها و مقادیر آنها:

ب)

سوال 2)

الف)

سناریو اول نشان دهنده overfitting است به دلیل تفاوت زیاد در درصد داده های آموزشی و تستی نشان میدهد مدل به خوبی با داده های آموزشی وفق دارد ولی در تست اینگونه نیست

سناریو دوم نشاندهنده Underfitting است چرا که هردو داده آموزشی و تست دارای دقت پایینی هستند

سناریو سوم نشان دهنده overfitting است به دلیل تفاوت زیاد در درصد داده های آموزشی و تستی نشان میدهد مدل به خوبی با داده های آموزشی وفق دارد ولی در تست اینگونه نیست و توانایی عمومی سازی ندارد

ب)

سناریوی 1 (بیش‌برازش)

استفاده از تکنیک‌های regularization مانند L1 و L2

کاهش تعداد نورون‌ها یا کاهش پیچیدگی مدل

استفاده از dropout برای کاهش وابستگی نورون‌ها به یکدیگر

سناریوی 2 (پیش‌برازش)

افزایش تعداد نورون‌ها در لایه پنهان برای افزایش ظرفیت مدل

استفاده از مدل‌های پیچیده‌تر

افزایش تعداد دوره‌های آموزشی (epochs) تا مدل به خوبی آموزش داده شود

سناریوی 3 (بیش‌برازش)

استفاده از تکنیک‌های regularization مانند L1 و L2

استفاده از dropout برای کاهش وابستگی نورون‌ها به یکدیگر

کاهش تعداد نورون‌ها در لایه پنهان یا استفاده از معماری‌های ساده‌تر

سوال 3)

الف و ب)

Dropout: روش برای جلوگیری از overfitting می باشد. وفتی یک لایه fully connected تعداد زیادی نورون دارد، شزایطی میتوناد پیش بیاید که در آن تعدادی از نورونی ها ویژگی های یکسان یا شبیه به هم را از داده ها استخراج میکنند که میتواند منجر به overfitting بشود. با استفاده از dropout درصدی از نورون ها به صورت تصادفی خاموش میشوند. این روش فقط در لایه های پنهان استفاده میشود

Regularization: یک تکنیک برای رفع overfitting است که در آن یک penalty به loss function اضافه میشود و مدل را از اهمیت دادن به بعضی از ویژگی ها و ضرایب برحضر میدارد

ج) تفاوت L1 و L2 در این است که L2 از مربع وزن ها استفاده میکند.

زمانی از L1 استفاده میکنیم که در آموزش مدل خود داریم feature selection انجام میدهیم و تنها میخواهیم از بخشی از feature های داده ها که مهم هستند استفاده کنیم

زمانی از L2 استفاده میکنیم که میخواهیم ضرایب را کاهش دهیم ولی همچنان به تمام featureها نیاز داریم و نمیخواهیم آنها زا حذف کنیم و feature ها به مربوط هستند (multicollinearity)

سوال 4)

الف) SOMها یه self organizing map ها نوعی شبکه عصبی هستند که در unsupervised learning و data visualization استفاده میشوند و هدف اصلی آنها کاهش ابعاد داده ها در عین نگهداشتن رابطه میان آنها است (در clustering استفاده میشود). شامل یک لایه ورودی و یک لایه خروجی است

ب) میتوان وزن ها را صورت تصادفی مفداردهی کرد که ساده ترین راهکار است ولی منجر به همگرایی آهسته تر میشود. میتوان به صورت خطی مقداردهی کرد که شروع ساختار یفاته تری خواهیم داشت و به همگرایی سریع تری منجر میشود. اهمیت وزن دهی:

1- در سرعت همگرایی تاثیر دارد

2- در پایداری نگاشت ها تاثیر دارد و میتواند از دوری از کردن از مینیمم محلی کمک کند

3- کیفیت خوشه بندی را تحت تاثیر قرار میدهد به طوری که به خوشه بندی بهتر و نمایش بهتری از داده ها منجر میشود

ج)اگر تعداد گره ها کم باشد نمیتواند به درستی روابط داده ها را بیابد و منجر به underfitting میشود. در مقابل اکر تعداد گره ها زیاد باشد جزئیات بیشتری را بدست میاورد ولی میتواند منجر به overfitting میشود. در کنار این موارد میتواند زمان train کردن تاثیر بگذارد که هرچه تعداد گره ها بیشتر آپدیت های بیشتری باید رخ بدهد و حافظه بیشتری استفاده میکند

سوال 5)

الف)

1. ساخت مراکز به صورت تصادفی
2. تعیین فاصله هر داده با مراکز
3. انتساب یک داده به نزدیک ترین مرکز
4. تعیین مرکز جدید با توجه به داده اضافه شده
5. تکرار مراحل 3 و 4

ب)

از نوع unsupervised هستند از آنجایی که داده ها label ندارند

ج)

پیش برازش (underfitting)

د)

میان خوشه ای(intra cluster) را کم و درون خوشه ای(inter cluster) را زیاد میکنیم. هرچه میان خوشه ای کمتر، شباهت و ربط داده های درون یک خوشه بیشتر و هرچه درون خوشه ای بیشتر، overlap خوشه ها کمتر و متمایز تر هستند

سوال 6)

شبکه عصبی convolution یک کلاس از شبکه های عصبی است که برای پردازش موثر داده های grid شکل مانند عکس ها بکار میرود. این شبکه دارای چندین لایه است که به ترتیب برروی پردازش داده کار میکنند:

1. لایه convolution که با استفاده از filter ها ویژگی های ورودی را استخراج میکند
2. لایه pooling که ورودی را downscale میکند تا پیچیدگی های محاسباتی را کاهش دهد
3. لایه fully connected که وظیفه پیشبینی را برعهده دارند و خروجی را تولید میکنند

ب)

کاربردهای آن بیشتر در image classification و image recognition است و مزیت های آن عبارتند از:

1. نیاز به supervision ندارند
2. استخراج خودکار featureها
3. دقت بالا و کاهش محاسبات
4. قابلیت هندل کردن datasetهای بزرگ

ج)

از لایه pooling برای downscale کردن خروجی لایه convolution و در نتیجه کاهش حجم محاسبات استفاده میشود و به دو صورت max pooling و average pooling صورت میگیرد

د)

از فیلترها برای استخراج ویژگی ها از dataset استفاده میشود و به صورت که فیلترها برروی داده ورودی به اصطلاح لغزانده (slide) میشوند و حاصل ضرب بخصی از ورودی که فیلتر برروی آن است با فیلتر، خروجی لایه convolution را تشکیل میدهد و در ادامه با backprop وزن های موحود در فیلترها آپدیت میشوند.

سوال 7)



