

پاییز ۱۳۹۸

یادگیری ماشین

تمرین اول

پرهام الوانی

۹۸۱۳۱۹۱۰

(الف) یادگیری نظارتی: در الگوریتم‌های یادگیری نظارتی داده‌های آموزش شامل پاسخ نهایی می‌باشند. به این پاسخ نهایی برچسب می‌گوییم.

(ب) یادگیری نیم نظارتی: این الگوریتم‌ها از داده‌هایی استفاده می‌کنند که تماماً دارای برچسب نیستند. برای مثال سرویس Google Photos از تعدادی عکس دارای برچسب افراد برای شناسایی افراد در سایر عکس‌ها استفاده می‌کند.

(پ) یادگیری بدون نظارت: الگوریتم‌ها سعی می‌کنند از داده‌های بدون برچسب استفاده می‌کند و سعی می‌کند بدون معلم آموزش پیدا کند.

(ت) یادگیری تقویتی: الگوریتم‌های یادگیری تقویتی با مشاهده محیط، انتخاب و اعمال کنش در جواب امتیاز یا پاداش دریافت می‌کند. سیستم سعی می‌کند با یادگیری بهترین استراتژی بیشترین جایزه را در طول زمان بدست آورد.

(ث) رگرسیون: در صورتی که تابعی که می‌خواهیم آن را یاد بگیریم یک تابع پیوسته باشد به آن مساله یک مساله‌ی رگرسیون می‌گوییم.

(ج) یادگیری برخط: در مساله‌ی یادگیری برخط شما سیستم را به صورت افزایشی با وارد کردن نمونه‌های داده آموزش می‌دهید.

(چ) یادگیری فعال: در یادگیری فعال داده‌ها از طریق پرس‌وجوهایی که انجام می‌دهیم به دست می‌آیند.

(ح) یادگیری انتقالی: در این نوع یادگیری تلاش می‌شود که از دانشی که برای یک مساله بدست آمده است برای حل یک مساله‌ی دیگر استفاده کرد.

(خ) دسته‌بندی: در این مسائل می‌خواهیم تابعی را یاد بگیریم که وظیفه‌ی آن مشخص می‌کند هر داده به چه گروهی تعلق دارد.

(د) یادگیری قیاسی: برای این نوع از یادگیری از منطق قیاسی استفاده می‌شود که در آن از یک یا چند گزاره نتیجه‌گیری می‌کنند.

(ذ) یادگیری استنتاجی: در این یادگیری از گزاره‌ها به عنوان منابعی برای تایید نتیجه استفاده می‌شود. منطق استنتاجی برخلاف منطق قیاسی ممکن است نتایجی حاصل کند که کاملاً صحیح نباشند. به عنوان مثال در این یادگیری از گزاره‌ای که برای یک نمونه برقرار است برای همه‌ی اعضا استفاده می‌شود.

(ر) درون‌یابی: درون‌یابی روش عددی می‌باشد که از آن برای پیدا کردن مقدار یک تابع درون یک بازه با داشتن مقدارهایی از تابع استفاده می‌شود.

(ز) برون‌یابی: برون‌یابی روشی است که در آن از مقدارهای مشخص تابع در یک بازه برای پیدا کردن مقدار یک تابع در بازه‌ای دیگر استفاده می‌شود.

(ژ) بیش‌برازش: در صورتی که یک مدل یادگیری بیش از حد به داده‌های آموزشی توجه کند به صورتی که ممکن است باعث افزایش خطا با داده‌های جدید شود بیش‌برازش رخ داده است.

۲. مسائل رگرسیون یک تابع پیوسته را یاد می‌گیرند در حالی که مسائل دسته‌بندی تابع گسسته‌ای را یاد می‌گیرند.

به طور مثال با بازه‌بندی برد تابع رگرسیون می‌توان از آن برای مسائل دسته‌بندی استفاده کرد و برای استفاده از تابع دسته‌بندی در مسائل رگرسیون به طور مثال می‌توان برای هر دسته یک عدد انتخاب شود که از آن به عنوان خروجی تابع استفاده کرد.

۳. همبستگی آماره‌ای است که بین -1 تا 1 قرار می‌گیرد و نشان می‌دهد دو متغیر به چه میزان نسبت به یکدیگر تغییر می‌کنند. رابطه‌ی ریاضی این آماره به شرح زیر است:

در صورتی که بین ویژگی‌ها همبستگی وجود داشته باشد می‌توان برای مهندسی ویژگی‌ها آن را حذف کرد و تعداد ویژگی‌ها را کاهش داد. به این ترتیب مدل حاصل ساده‌تر می‌شود.

۴. در روش رگرسیون هدف انتخاب ضرایبی است که خطا را به صورت کلی روی مجموعه‌های داده‌های آموزش کمترین مقدار سازد ولی در روش *Locally Weighted Regression* به ضرایب وزن داده می‌شود که وزن به گونه‌ای است که خطا برای داده‌هایی که به نقطه‌ی داده شده نزدیکتر هستند تاثیر خطا بیشتر باشد. این روش ممکن است برای هر داده‌ی ورودی جدید مجموعه‌ی جدیدی از ضرایب را استفاده کند که این امر در رگرسیون وجود ندارد.

۵. در ادامه این دو روش بر پایه‌ی مساله‌ی رگرسیون خطی مقایسه می‌شوند:

روش گرادیان نزولی یک روش تکراری است که می‌تواند بعد از گام‌هایی می‌تواند به نقطه‌ی بهینه همگرا شود. روش معادله‌ی نرمال نیاز به محاسبه‌ی ماتریس معکوس دارد ولی از سوی دیگر در یک گام جواب بهینه را تولید می‌کند.

۶. الف) مدل بایاس با تغییر ویژگی‌ها تغییر اندکی می‌کند و این در حالی است که مدل واریانس با تغییر یک ویژگی به سرعت تغییر می‌کند. مدل واریانس روی داده‌های آموزش خطای کمی دارد این در حالی است که مدل بایاس می‌تواند خطای بیشتری داشته باشد.

ب) انتخاب این دو باید به گونه‌ای باشد که پیچیدگی مدل زیاد نباشد اما بتواند خطای کمی در عمل تولید کند.

پ) مدل واریانس در MSE مقدار کمی دارد ولی مدل بایاس خطای زیادتری تولید می‌کند.

۷. برای جلوگیری از *Overfit* شدن مدل، عمل *Regularization* صورت می‌گیرد. در این روش سعی می‌شود اهمیت جمله‌هایی که درجه‌ی بالا دارند کمتر شود. البته این امر به چگونگی تعریف رابطه‌ی *Regularization* نیز بستگی دارد.

۸. داده‌های ارزیابی برای انتخاب مدل استفاده می‌شوند و ممکن است چندبار استفاده شوند. در ادامه مدل با داده‌های آموزش، آموزش داده می‌شود. در نهایت از داده‌های آزمون برای ارزیابی مدل استفاده می‌شود. داده‌های آزمون می‌بایست کم استفاده شوند چرا که استفاده‌ی زیاد از آن‌ها می‌تواند مدل را بایاس کند.

۹. هر سه این فرمول‌ها هدف تعریف خطا را دارند. در هر یک از این روش‌ها اهمیتی که به خطا داده می‌شود متفاوت است. در صورتی که از MAE استفاده شود خطا با درجه ۱، در صورتی که از MSE استفاده خطا با درجه ۲ به دست می‌آید. در RMSE در نهایت از مجموع مربعات خطا جذر گرفته می‌شود.