



پردازش هوشمند تصاویر زیست پزشکی

نیم سال اول ۰۳-۰۲

مدرس: محمدحسین رهبان

دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

مدت زمان: ۲۵ دقیقه

کوئیز ششم (۵ نمره)

۱. توزیع احتمال وابسته به زمان زیر را در نظر بگیرید:

$$p(x, t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp \left\{ -\frac{(x - x_0)^2}{2\sigma^2 t} \right\}.$$

تابع امتیاز $s(x, t)$ را برای این توزیع بدست آورید. (۱ نمره)
پاسخ:

$$s(x, t) = -\frac{(x - x_0)^2}{2\sigma^2 t}$$

۲. مدل آموزش دیده‌ای داریم که می‌خواهیم بر روی آن تحلیل خطا انجام دهیم. برای این کار ۱۰۰ مورد از تصاویر موجود در dev set را که مدل آن‌ها را اشتباه پیشبینی کرده است را بررسی می‌کنیم و خطاهای این تصاویر را در جدولی مشابه جدول زیر وارد می‌کنیم. با توجه به این جدول به سوالات زیر پاسخ دهید:

Image	Dog	Great cat	Blurry	Mislabeled	Comments
...					
98				✓	Labeler missed cat in background
99		✓			
100				✓	Drawing of a cat; not a real cat.
% of total	8%	43%	61%	6%	

۱. با توجه به اعداد جدول به نظر شما آیا تمرکز بر روی تمیز کردن داده‌های mislabeled در دیتاست ضروری است یا خیر؟ به صورت مختصر توضیح دهید. (۱ نمره)

پاسخ:

همانطور که در جدول بالا دیده می‌شود تنها ۶ مورد از ۱۰۰ خطای شبکه به خاطر اشتباه در برچسب زنی بوده است. بنابراین تمرکز بر روی این بخش اولویت کمتری نسبت به قسمت‌های دیگر دارد. ضمن اینکه برچسب زنی دوباره کل دیتاست زحمت زیادی دارد در این مرحله بهتر است به خطاهای دیگر پرداخت.

۲. به نظر شما اگر بخواهیم بهترین نتیجه را در آموزش دوباره مدل بگیریم، تمرکز بر روی بهبود کیفیت در کدام جنبه از مدل مهم‌تر است و اولویت بالاتری دارد؟ مختصر توضیح دهید. (۱ نمره)

پاسخ:

در حالت کلی در هر ایتريشن باید به خطاهایی پرداخت که هم اصلاح آنها آسان تر باشد و هم تاثیر زیادی در عملکرد کل مدل داشته باشد. با توجه به جدول بالا ۶۱ مورد از ۱۰۰ مورد خطا به خاطر تصاویر blurry بوده است. بنابراین در این مرحله بهتر است بر روی این مورد تمرکز شود که در صورت حل آن بهبود زیادی در عملکرد کلی مدل حاصل خواهد شد.

۳. یک مجموعه داده تصویر CT از ریه بیماران وجود دارد. این مجموعه توسط یک دانشجوی پزشکی برجسب سالم/بیمار خورده است. ما برای اطمینان از این که این برجسب ها خوب بوده یا نه، این مجموعه داده را به یک متخصص بسیار با تجربه بیماری های ریوی نشان دادیم و او نیز برجسب سالم/بیمار را تخصیص داد. در نتیجه مقایسه این دو برجسب گذاری مشاهده کردیم که آن دانشجوی پزشکی در ۵ درصد از دادگان اشتباه برجسب گذاری کرده است (فرض کنید که متخصصی که ما به آن مراجعه کردیم هیچ وقت اشتباه نمی کند). ما از این مجموعه داده برای آموزش یک مدل دسته بند نیز استفاده کردیم. بر این اساس به سوالات زیر پاسخ دهید:

۱. اگر خطای بخش آموزش برابر ۳۰ درصد شود و خطای بخش اعتبارسنجی (dev error) برابر ۳۵ درصد؛ به نظر شما مشکلی که پیش آمده است چیست؟ راه حلی پیشنهاد دهید. (۱ نمره)

پاسخ:

مشکل Underfit اتفاق افتاده است. یعنی مدل یادگیری ماشین استفاده شده یا آنقدر آموزش ندیده است که بتواند به خوبی کلاس ها را از هم تفکیک کند و یا آنقدر پیچیده نیست که بتواند فضا را مدل کند. چون اختلاف بین توانایی انسان در دسته بندی و مدل روی دادگان آموزش بسیار زیاد بوده است. به عنوان یک راهکار می توان از مدل پیچیده تری برای آموزش استفاده کرد.

۲. اگر خطای بخش آموزش برابر ۶ درصد شود ولی خطای بخش اعتبارسنجی (dev error) برابر ۴۰ درصد، به نظر شما مشکل چیست؟ راه حلی پیشنهاد دهید. (۱ نمره)

پاسخ:

مشکل Overfit اتفاق افتاده است. چون مدل توانسته است روی دادگان آموزش به دقتی تقریباً برابر انسان برسد ولی اختلاف جدی ای با بخش اعتبارسنجی دارد که نشان از بیش برآزش روی دادگان آموزش است. به عنوان یک راهکار می توان از منظم ساز بر روی پارامترهای مدل استفاده کرد.