

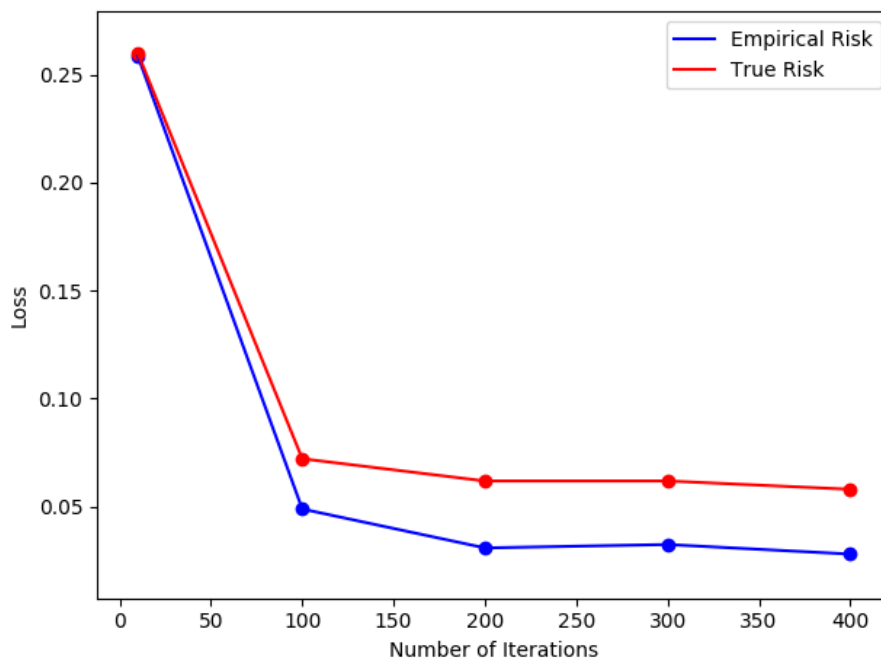
به نام خدا

گزارش تمرین سری ششم مقدمه‌ای بر یادگیری ماشین

حسین ابراهیمی - ۹۵۱۰۵۳۰۲

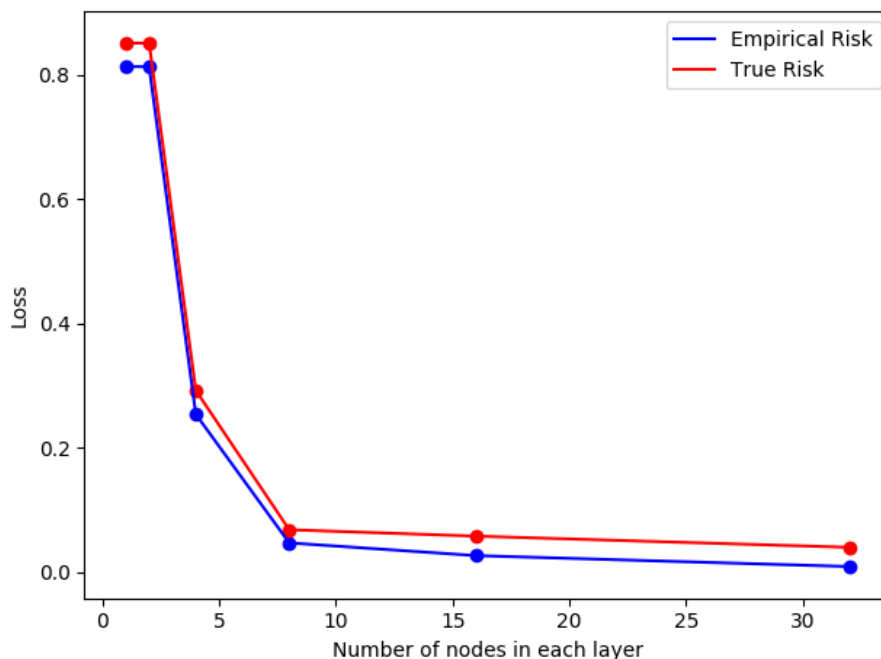
مدرس: دکتر گلستانی

الف.



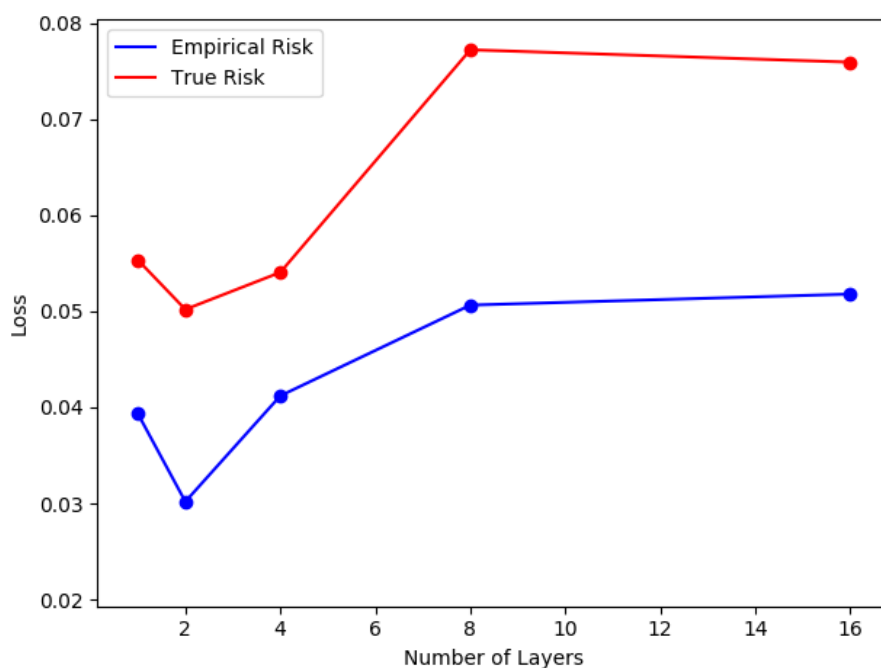
هر دو خطا در مرحله پنجم که تعداد گام‌ها برابر با ۴۰۰ است، کمترین مقدار خود را دارند. در ابتدا وقتی از ۱۰ گام به ۱۰۰ گام می‌رویم خطا به طور شدیدی کاهش پیدا می‌کند چون الگوریتم می‌تواند در فرصت بیشتری وزن‌های خود را بهبود ببخشد ولی از مرحله سوم به بعد یعنی برای ۲۰۰ گام و بیشتر الگوریتم تقریباً همگرا می‌شود و نوسان‌های کم‌ای از خود نشان می‌دهد که به دلیل stochastic optimizer هست و اگر batch size دهیم مقدار این نوسانات کاهش می‌یابد.

ب.



با توجه به اینکه feature size ما نسبتاً بزرگ است و ۵۶۱ ویژگی داریم، این باعث می‌شود که complexity مسئله و یا به عبارتی VC dim ورودی نیز بزرگ باشد. حال ما با اضافه کردن تعداد گره‌های بیشتر در هر لایه به ظرفیت شبکه‌ی عصبی می‌افزاییم و این سبب می‌شود که coding بهتری در هر لایه داشته باشیم و در نتیجه performance شبکه عصبی بهبود می‌یابد. در نتیجه با افزایش تعداد گره‌ها خطای تجربی و خطای واقعی هر دو کاهش می‌یابند.

ج.



کمترین خطا زمانی اتفاق می افتد که از دو hidden layer استفاده کنیم. به نظر می رسد وقتی تعداد لایه ها را افزایش دهیم چون ظرفیت شبکه عصبی افزایش می یابد، باید کارایی آن نیز متناسب با تعداد لایه بهبود یابد ولی اینطور نیست. بعد از کاهش خطا در اولین تغییر، روند خطای تجربی صعودی می شود و انگار مدل نامناسبی هر دفعه آموزش داده می شود و متناسب با آن خطای واقعی روی داده تست نیز بیشتر می شود. از طرف دیگر می توان اینطور بررسی کرد که داده مورد نظر با مدل های ساده ای fit می شود و با اضافه کردن تعداد لایه و در نتیجه افزایش پیچیدگی شبکه عصبی به مدل غلطی منجر می شود که هر دو خطای آموزشی و واقعی را افزایش می دهد.