

شبکه‌های عصبی و یادگیری عمیق

دکتر صفابخش



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر

رضا آدینه پور ۴۰۲۱۳۱۰۵۵

تمرین پنجم
شبکه‌های RNN

۱۰ خرداد ۱۴۰۳

سوال اول - نظری

به سوالات زیر بصورت خلاصه و برای هر یک حداکثر در سه بند پاسخ دهید:

۱. به‌طور کلی بهینه‌سازها^۱ (نظیر ADAM) به دنبال یافتن وزن‌های شبکه‌های عصبی هستند بطوریکه توابع هزینه^۲ کمینه شود. مشتق‌پذیر بودن توابع یاد شده چه تاثیری در بهینه‌ساز دارد؟ اگر مشتق‌پذیر نباشد، چه رویکردهایی برای بهینه‌سازی آن وجود دارد؟ یک مورد را به دلخواه توضیح دهید.
۲. محدب^۳ بودن توابع به چه معناست و چرا مطلوب است که در بهینه‌سازی، توابع هزینه محدب باشد؟ اگر محدب نباشد، چگونه می‌توان آن را بهینه نمود؟
۳. الگوریتم بهینه‌سازی نیوتن را مطالعه کرده و آن را با نزول در راستای گرادیان^۴ مقایسه کنید. در چه نوع مسائلی استفاده از الگوریتم نیوتن ارجحیت دارد؟
۴. ضمن مطالعه کلی الگوریتم AdaGrad، بیان کنید که چگونه می‌توان از آن برای بهینه ساختن نرخ یادگیری^۵ بهره گرفت. فرض کنید مسئله‌ی دسته بندی دودویی بحرانی بودن/نبودن شرایط یک کارگاه صنعتی بر اساس اطلاعاتی محیطی آن را در اختیار دارید که داده‌های دما، رطوبت، فشار و ذرات معلق بر اساس سنسورهای نصب شده در هر یک ثانیه ارسال می‌گردد. شما بایستی با در نظر گرفتن دنباله‌ای از داده‌های ارسالی بتوانید تشخیص دهید که شرایط بحرانی است یا خیر.

^۱Optimizer

^۲Loss Functions

^۳Convex

^۴Gradient Descent

^۵Learning Rate