

به نام خدا

محسن رحیمی

بخش اول:

Cross-validation یک تکنیک مهم در ارزیابی عملکرد مدل‌های یادگیری ماشین است که به منظور افزایش قابلیت اعتماد و کاهش احتمال برآزش بیش از حد (**overfitting**) مدل به داده‌های آموزش، استفاده می‌شود. این تکنیک به اجزای داده‌ها به نام "**fold**" تقسیم می‌شود و مدل روی این اجزا آموزش داده می‌شود و عملکرد آن روی بخش باقی‌مانده از داده‌ها ارزیابی می‌شود. این فرآیند چندین بار تکرار می‌شود، و در هر مرحله، یک **fold** به عنوان داده آزمون (**validation set**) و بقیه به عنوان داده آموزش (**training set**) استفاده می‌شود.

مزایای استفاده از **cross-validation** عبارتند از:

1. **کاهش وابستگی به تقسیم‌بندی داده:** استفاده از یک تقسیم‌بندی خاص از داده ممکن است موجب ایجاد یک مدل خاص به آن تقسیم‌بندی شود. اما با استفاده از **cross-validation**، ما از چندین تقسیم‌بندی مختلف برای آموزش و ارزیابی مدل استفاده می‌کنیم.
2. **افزایش دقت ارزیابی:** میانگین عملکرد مدل بر روی تمام **folds** به عنوان یک معیار قوی‌تر برای ارزیابی مدل در مقایسه با استفاده از یک بار تقسیم‌بندی مشخص است. این باعث می‌شود که نتایج معتبرتر و قابل اعتمادتری برای عملکرد مدل در مقابل داده‌های جدید داشته باشیم.
3. **کاهش اثرات تصادفی:** زمانی که داده‌ها به صورت تصادفی تقسیم می‌شوند، ممکن است یک تقسیم‌بندی خاص به دلیل ویژگی‌های خاص داده‌ها به مدل موجب افزایش دقت نتایج شود. استفاده از چندین **folds** باعث کاهش تاثیر این تصادفات ممکن می‌شود.

هر یک از **fold** ها به ترتیب به عنوان داده آزمون استفاده می‌شوند و مدل روی بقیه **folds** آموزش داده می‌شود. این فرآیند چندین بار تکرار می‌شود و میانگین عملکرد مدل در همه **folds** به عنوان یک معیار کلی برای ارزیابی مدل در نهایت گزارش می‌شود.

بخش دوم:

در این بخش نمودار توابع مختلف پیش بینی شده است. برای افزایش دقت پیش بینی از تکنیک‌هایی از جمله افزایش تراکم نقطه‌ها، افزایش تعداد نوروں‌های هر لایه و افزایش تعداد لایه‌ها استفاده شده است که موثرترین روش افزایش تعداد نمونه‌ها بوده.

در انتهای این بخش نمودار های ۳ بعدی نیز به شبکه عصبی داده شده و با دقت خوبی تقریب زده شده اند.

بخش سوم:

در این بخش یک نویز نرمال به داده های آموزشی اضافه شده و شبکه روی آن فیت شده است و نتیجه نمایش داده شده است.

بخش چهارم:

در این بخش نمودار دستی وارد شده است و با شبکه عصبی تقریب زده شده که در ابتدا جزئیات تابع به درستی تقریب زده نشد ولی با افزایش تعداد نمونه ها مشکل حل شد و دقت خوبی به دست آمد. برای افزایش تعداد نقاط از درونیابی خطی استفاده شده به طوری که بین هر دو نقطه یک خط رسم و تعدادی نقطه روی آن اضافه شده است.

بخش پنجم:

در این بخش دیتا های شماره ها به شبکه عصبی داده شد و با یک شبکه کوچک دقت ۹۴ درصد به دست آمد. با بزرگ تر کردن شبکه دقت تا ۹۶ درصد هم افزایش یافت. بعد برای جلوگیری از overfitting از cross-validation با تعداد fold های ۵ استفاده شده و به دقت میانگین ۹۶ درصد رسیدیم.

بخش ششم:

در این بخش ابتدا یک نویز نرمال به تصاویر اضافه شده و تصاویر در یک مسیر ذخیره شدند. در ادامه از شبکه عصبی با activation function لاجستیک استفاده شد تا نویز ها حذف شوند. و در بهترین حالت فاکتور psnr به ۲۰ رسید.