

## تمرین سری چهارم

درس: یادگیری ماشین- پاییز ۱۴۰۳

استاد درس: دکتر فاطمه زمانی

دستیاران آموزشی: حسین آقاگل زاده، ابوالفضل حسینی فر

دانشکده مهندسی برق و کامپیوتر دانشگاه صنعتی نوشیروانی بابل

- این تمرین دارای سه بخش تئوری، پیاده سازی و مطالعاتی است.
- پاسخنامه بخش تئوری می تواند به صورت دست نویس یا تایپی باشد. در هر صورت اما می بایست در قالب یک فایل pdf تحویل گردد.
- پاسخ بخش پیاده سازی می بایست حاوی کد پیاده سازی شده به همراه گزارش آن در قالب فایل pdf باشد. از آوردن کد در گزارش اجتناب کنید (مگر آوردن بخشی از آن از نظر شما ضرورت داشته باشد).
- پاسخ بخش مطالعاتی می بایست به صورت تایپ شده در قالب یک فایل pdf باشد.
- گزارش بخش پیاده سازی می بایست میزان تلاش شما رو با ۴ اولویت مهم ۱- شفافیت ۲- درستی ۳-زیبایی ۴- کوتاهی، نشان بده
- بخش تئوری بدون استفاده از کد انجام بشه مگر دلیل موجهی وجود داشته باشه.
- کل پاسخ (کد و فایل های pdf) در یک فایل zip با فرمت زیر ارسال شود.

HW4.zip شماره دانشجویی\_

## تئوری

### مسئله ۱

در روش mixture of experts می توان از expert هایی که ورودی representation های متفاوتی دارند استفاده شود. در این سناریو پیشنهاد شما برای طراحی شبکه gating چیست؟

### مسئله ۲

فرض کنید دنبال به کارگیری یک ساختار mixture of experts با رای گیری وزن یکسان و با ورودی های کاملاً یکسان، شامل ۳ مدل مستقل هستیم. در نظر بگیرید به ما اجازه داده شده تا ۵ مدل را آموزش بدیم ولی در نهایت حق استفاده از فقط ۳ مدل از آن ها را داریم. نتیجه ارزیابی بر روی دیتای validation بر روی ۵ مدل پس از آموزش به صورت زیر حاصل شده. تعداد داده های validation ۶ عدد است و مقدار ۱ به معنی تشخیص درست و ۰-۱ به معنی تشخیص نادرست است. از نظر شما کدام مدل ها باید انتخاب شوند؟

$$\hat{y}_1 : (-1, -1, 1, -1, 1, -1)$$

$$\hat{y}_2 : (1, 1, 1, -1, 1, -1)$$

$$\hat{y}_3 : (1, -1, -1, 1, 1, 1)$$

$$\hat{y}_4 : (1, 1, 1, -1, -1, -1)$$

$$\hat{y}_5 : (-1, 1, 1, 1, -1, 1)$$

### مسئله ۳

در یادگیری ماشین، در نگاهی از خطای مدلی که قصد آموزش آن را داریم دو خطای اساسی وجود دارد. Bias و Variance.

الف) این دو نوع خطا را با یکدیگر مقایسه کنید.

ب) در مورد ارتباط آن ها با مسئله underfitting/overfitting بحث کنید.

پ) به نظر شما هدف اصلی بحث ترکیب classifier ها کاهش کدام یک از این دو خطا است؟ توضیح دهید.

## مسئله ۴

الف) دو روش تقسیم بندی دیتای K-Fold و bootstrap را به همراه تفاوت هایشان توضیح دهید.

ب) در روش bagging استفاده از کدام یک را پیشنهاد می دهید؟ چرا؟

## پیاده سازی

### مسئله ۱

دیتاست MNIST را به برنامه اضافه کنید و آن را به دو بخش train (۶۰۰۰۰ نمونه) و test (۱۰۰۰۰ نمونه) تقسیم کنید.

- به جز برای درخت تصمیم مجاز به استفاده از دستور آماده نیستید (مگر خلافتش ذکر شود).

**الف)** با کمک کتابخانه sklearn با روش درخت تصمیم عملیات آموزش و ارزیابی را بر روی دیتاست انجام دهید.

**ب)** ۲۰ درخت تصمیم مستقل در نظر بگیرید و هر بار یک درخت را با ۵۰ درصد از دیتای آموزشی که به طور تصادفی انتخاب میشود و شامل ۷۰ درصد ویژگی ها که باز هم تصادفی انتخاب می شوند آموزش دهید.

\* سراغ هر درخت تصمیم بعدی که می روید انتخاب ۵۰ درصد دیتا و ۷۰ درصد ویژگی ها را از اول تصادفی انجام دهید.

**پ)** مرحله آزمون را ۲۰ درخت تصمیمی که به صورتی ترکیبی با دو سیستم voting زیر کار می کنند انجام دهید.

- بر اساس رای اکثریت

- سیستم وزن دهی به صورتی که درخت تصمیمی که دقت بیشتری روی دیتای آموزشی داشته است وزن بیشتری داشته باشد. راهنمایی: دقت های حاصل از ۲۰ درخت بر روی دیتای train را به جمع دقت ها تقسیم کنید (نرمال سازی).

**ت)** در این مسئله سعی در پیاده سازی الگوریتم random forest که یک روش ترکیب classifier ها است داشتیم. حال این بار با کمک کتابخانه sklearn مراحل آموزش و آزمون را بر روی دیتا انجام داده و نتایج بدست آمده را با قسمت پ مقایسه کنید.

### مسئله ۲

- به جز برای درخت تصمیم مجاز به استفاده از دستور آماده نیستید (مگر خلافتش ذکر شود).

روش ترکیبی Adaboost را با ۲۰ درخت تصمیم بر روی MNIST انجام دهید و نتیجه را با نتایج استفاده از دستور آماده آن در کتابخانه sklearn مقایسه کنید.

## مطالعاتی

مقاله ی پیوست شده را مطالعه کنید و برداشت خود از آن را شرح دهید (حداقل ۲ صفحه).

موفق باشید 😊