



دانشگاه اصفهان - رشته علوم کامپیوتر  
یادگیری ماشین

## تمرین ۲: بیش‌برازش و راه‌های جلوگیری از آن

شماره دانشجویی:

نام و نام خانوادگی:

### پرسش ۱.

بیش‌برازش<sup>۱</sup> و کم‌برازش<sup>۲</sup> را توضیح دهید.

پاسخ.

### پرسش ۲.

فرض کنید  $H$  فضای فرضیه در نظر گرفته شده برای یک مسئله یا مفهوم<sup>۳</sup> است. از میان تمام فرضیه‌هایی که درون مجموعه  $H$  قرار دارند، کدام یک بهترین فرضیه برای یک مسئله است؟ آیا الزاماً این فرضیه دارای خطای تجربی<sup>۴</sup> صفر است؟ توضیح دهید.

پاسخ.

### پرسش ۳.

الف. می‌دانیم که یک راه برای شناسایی بیش‌برازش و کم‌برازش، استفاده از مجموعه داده اعتبارسنجی<sup>۵</sup> است. این مجموعه داده از جنس کدام یک از مجموعه‌های آموزش (داده دیده شده توسط مدل) و آزمایش (داده ندیده شده) است؟ علت خود را شرح دهید.

ب. آیا ممکن است با استفاده از مجموعه داده اعتبارسنجی بتوانیم بیش‌برازش یا کم‌برازش را تشخیص بدهیم؟ توضیح دهید.

<sup>1</sup>Overfitting

<sup>2</sup>Underfitting

<sup>3</sup>Concept

<sup>4</sup>Empirical Error

<sup>5</sup>Validation Set

پاسخ.

الف.

ب.

پرسش ۴ (اختیاری).

الف. توضیح دهید روش منظم‌سازی<sup>۶</sup> چگونه باعث جلوگیری از بیش‌برازش و کم‌برازش می‌شود.

ب. یکی از روش‌های  $L2$  یا  $L1$  را به دلخواه انتخاب کرده و رابطه آن را از طریق تغییر فرم بهینه‌سازی تابع هزینه بدست آوردید.

پاسخ.

الف.

ب.

---

<sup>6</sup>Regularization