

پاسخنامه

پاسخنامه کونیز جلسه 10 دوره علم داده، یادگیری ماشین و هوش مصنوعی - فرزاد مینویی

✓* ۱- داده‌های یک نیروگاه تولید برق در بازه 6 ساله در قالب داده‌های آموزش و آزمایش در فایل‌های hw10_train.csv و hw10_test.csv ذخیره شده است. این فایل‌ها در گوگل درایو درس در پوشه S10 موجود است. داده‌ها مربوط به زمانی است که نیروگاه برای کار با حداکثر ظرفیت تنظیم شده است.

هدف پیش‌بینی انرژی الکتریکی خالص ساعتی (PE) بر حسب مگاوات ساعت است. برای این منظور متوسط ساعتی دمای محیط (AT) بر حسب سانتی‌گراد، فشار محیط (AP) بر حسب سانتی‌متر جیوه، رطوبت نسبی (RH) بر حسب درصد و فشار گاز خروجی از توربین (V) بر حسب سانتی‌متر جیوه بعنوان متغیرهای پیش‌بینی‌کننده در نظر گرفته شده‌اند.

ابتدا تحلیل اکتشافی روی داده‌های آموزش انجام دهید. کدام گزینه درست است؟

۱ برخی متغیرهای این مسئله از نوع رسته‌ای هستند.

۲ میانگین متغیر پاسخ 468 است.

۳ متغیر رطوبت نسبی چولگی به راست دارد.

✓ ۴ ضریب همبستگی پیرسون بین متغیر پاسخ و دمای محیط 95- درصد است.

۲* - روی داده‌های آموزش hw10_train.csv یک مدل نزدیکترین همسایه (kNN) بسازید. تعداد همسایه‌ها را 25 قرار دهید. شاخص MSE روی داده‌های آموزش چقدر است؟

3.5 ۱

✓ 16.7 ۲

1.6 ۳

4.1 ۴

۳* - روی داده‌های آموزش hw10_train.csv یک مدل نزدیکترین همسایه (kNN) بسازید. تعداد همسایه‌ها را 25 قرار دهید. شاخص RMSE روی داده‌های آموزش چقدر است؟

✓ 4.1 ۱

1.5 ۲

3.2 ۳

16.7 ۴

✓*۴- روی داده‌های آموزش hw10_train.csv یک مدل نزدیکترین همسایه (kNN) بسازید. تعداد همسایه‌ها را 25 قرار دهید. شاخص MAE روی داده‌های آموزش چقدر است؟

۱ 6.7

۲ 4.1

✓ ۳ 3.1

۴ 2.8

✓*۵- روی داده‌های آموزش hw10_train.csv یک مدل نزدیکترین همسایه (kNN) بسازید. تعداد همسایه‌ها را 25 قرار دهید. شاخص MAPE روی داده‌های آموزش چند درصد است؟

✓ ۱ 0.7

۲ 1.2

۳ 2.3

۴ 7

✓ ۶- روی داده‌های آموزش hw10_train.csv یک مدل نزدیکترین همسایه (kNN) بسازید. تعداد همسایه‌ها را 50 قرار دهید. شاخص RMSE روی داده‌های آموزش چقدر است؟

- ✓ ۱ 4.4
- ۲ 5.6
- ۳ 8.3
- ۴ 7.1

✓ ۷- فرآپارامترهای زیر را برای الگوریتم نزدیکترین همسایه در نظر بگیرید.

[50, 45, 40, 35, 30, 25, 20, 15, 10, 5, 1]

به منظور پیدا کردن بهترین عدد برای فرآپارامتر مدل، الگوریتم k-Fold Cross Validation را بکار بگیرید. تعداد پوشه‌ها را 10 در نظر بگیرید. مقدار بهینه فرآپارامتر چقدر است؟

- ✓ ۱ 5
- ۲ 15
- ۳ 25
- ۴ 40

۸* ✓ - فرآپارامترهای زیر را برای الگوریتم نزدیکترین همسایه در نظر بگیرید.

[50,45,40,35,30,25,20,15,10,5,1]

به منظور پیدا کردن بهترین عدد برای فرآپارامتر مدل، الگوریتم k-Fold Cross Validation را بکار بگیرید. تعداد پوشه‌ها را 10 در نظر بگیرید. مقدار متوسط شاخص RMSE برای حالت بهینه چقدر است؟

۱ 4.49

۲ 4.25

✓ ۳ 4.01 ✓

۴ 4.17

۹* ✓ - مدل با بهترین فرآپارامتر را روی داده hw10_train.csv، آموزش دهید. سپس برای ارزیابی نهایی، از آن برای پیش‌بینی روی داده آزمایش (hw10_test.csv) استفاده کنید. مقدار شاخص RMSE روی داده آزمایش چقدر است؟

✓ ۱ 4.1 ✓

۲ 7.2

۳ 16.7

۴ 19.2

✓ ۱۰- اگر تحت شرایط جدیدی، متوسط ساعتی دمای محیط 22 درجه سانتی‌گراد، فشار محیط 1010.5 سانتی‌متر جیوه، رطوبت نسبی 75 درصد و فشار گاز خروجی از توربین 68.5 سانتی‌متر جیوه باشد، از مدل نهایی استفاده کنید و برآورد انرژی الکتریکی خالص ساعتی را پیش‌بینی کنید.

۱ 421.3

۲ 435.8

✓ ۳ 441.5

۴ 475.5

Porsline

با پرس‌لاین فرم خود را طراحی و ارسال کنید و گزارش‌های آماری دریافت نمایید.