

## گزارش نهایی

هدف تمرین، پیاده سازی الگوریتم درخت تصمیم برای تحلیل و پیشبینی داده ها بوده است. درخت تصمیم یک مدل یادگیری نظارت شده است که به کمک مجموعه ای از ویژگی ها تصمیم گیری میکند. این مدل عمدتاً در مسائل دسته بندی و رگرسیون استفاده میشود و با ساختار درختی خود، جریان تصمیم گیری را نشان میدهد.

### • پیش پردازش ها

قبل از آموزش مدل، مجموعه داده مورد بررسی قرار گرفت و مراحل زیر روی آن انجام شد:

- **بررسی داده های ناقص:** داده های نامعتبر حذف یا با مقادیر مناسب جایگزین شدند.
- **نرمال سازی داده ها:** در صورت نیاز، داده ها استاندارد شدند تا تأثیر ویژگی ها به صورت یکنواخت باشد.
- **تقسیم داده ها:** مجموعه داده به دو بخش آموزش (۷۰٪) و آزمون (۳۰٪) تقسیم شد تا عملکرد مدل روی داده های دیده نشده ارزیابی شود.

### • تحلیل نتایج

برای بررسی کیفیت مدل، معیارهای ارزیابی محاسبه شدند:

- **Accuracy:** میزان دقت کلی مدل در پیشبینیها.
  - **Precision:** میزان دقت در تشخیص کلاسهای مثبت.
  - **Recall:** توانایی مدل در شناسایی موارد مثبت.
  - **F1-Score:** میانگین ترکیبی Precision و Recall.
- همچنین نمودار درخت تصمیم رسم شد تا مسیرهای تصمیم گیری مدل بررسی شود. مشاهده شد که مدل بر اساس ویژگی های انتخاب شده، شاخه بندی انجام داده و تصمیمات طبقه بندی را اتخاذ میکند.

### • نقاط قوت و ضعف مدل

#### ✓ نقاط قوت:

- تفسیرپذیری بالا؛ مدل به صورت بصری قابل درک است.
- امکان بررسی دقیق تأثیر ویژگیها بر تصمیمات مدل.

#### ✗ نقاط ضعف:

- احتمال بیشبرازش (Overfitting)؛ اگر درخت بیشازحد پیچیده شود، عملکرد مدل روی دادههای جدید کاهش مییابد.
- حساسیت به دادههای پرت، که میتواند تصمیمات مدل را تحت تأثیر قرار دهد.

### • نتیجه گیری

مدل درخت تصمیم ابزار قدرتمندی برای تحلیل داده ها و پیش بینی کلاس ها است. در این تمرین، با پردازش داده ها و آموزش مدل، عملکرد درخت تصمیم بررسی شد. با توجه به تحلیل نتایج، پیشنهاد میشود هرس کردن درخت (Pruning) و تنظیم پارامترهای مدل برای جلوگیری از بیش برآزش مورد استفاده قرار گیرد.