

## تمرین سه

### شبکه عصبی

## ۱ تشخیص تقلب

این مجموعه داده متشکل از اطلاعات تراکنش‌های مالی است که با استفاده از این داده‌ها و یک مدل AutoEncoder باید جعلی بودن یک تراکنش تشخیص داده شود. حتما در گزارش تحلیل کافی بر روی عملکرد مدل و دلیل کارایی آن توضیح داده شود.

ساختار مجموعه داده به صورت زیر است. ستون isFraud نشان‌دهنده جعلی یا غیر جعلی بودن تراکنش است.

### ● جدول تراکنش‌ها

- TransactionDT: اختلاف زمانی انجام شدن تراکنش از یک مرجع زمانی.
- TransactionAMT: مبلغ تراکنش به واحد دلار آمریکا.
- ProductCD: کد محصول یا سرویس ارائه شده در تراکنش.
- card1 - card6: اطلاعات کارت مورد استفاده برای پرداخت مثل نوع کارت، دسته کارت، بانک مرجع، کشور و ...
- addr: کشور و منطقه پرداخت.
- dist: فاصله بین آدرس پرداخت، آدرس پستی، کد پستی، آدرس، IP، منطقه شماره تلفن و ...
- emaildomain (P\_ and R\_): ایمیل خریدار و فروشنده.
- C1-C14: داده از جنس شمارنده، مثل تعداد آدرس‌های مربوط به یک کارت پرداخت و... (معنای اصلی این ستون‌ها مخفی شده است).
- D1-D15: اختلاف زمانی، مثل تعداد روزها نسبت به تراکنش قبلی.
- M1-M9: داده از جنس ارتباطات، مثل ارتباط نام روی کارت و آدرس و...
- Vxxx: ویژگی‌های غنی ساخته شده، مثل رتبه، تعداد و دیگر ارتباطات موجودیت‌ها.

### ● جدول شناسه‌ها

- متغیرهای این جدول اطلاعات شناسه‌ای مانند اطلاعات اتصال به شبکه (IP, ISP, Proxy, etc) و امضا الکترونیکی (UA/browser/os/version, etc) مربوط به تراکنش است. نام اصلی این ویژگی‌ها مخفی شده است.