



دانشگاه مازندران
گروه علوم کامپیوتر

پروژه پایانی - درس یادگیری ماشین

عنوان پروژه: کشف تقلب در تراکنش های مالی با استفاده از الگوریتم های یادگیری ماشین

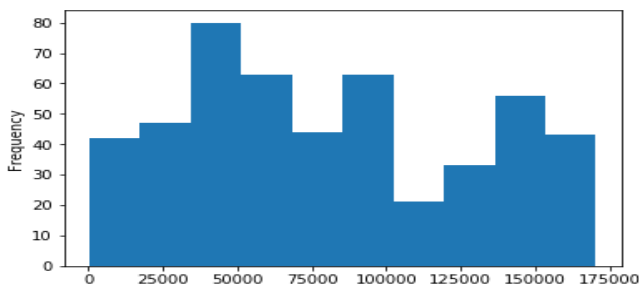
آخرین مهلت تحویل پروژه: ۱۴۰۲/۰۴/۲۵ | نمره پروژه: ۱۲ نمره

هدف پروژه

کشف تقلب در تراکنش های مالی یکی از اهداف حیاتی در بانک ها و موسسات صدور کننده کارت های اعتباری است. هدف از این پروژه استفاده از یک مجموعه داده شامل اطلاعات تعداد زیادی تراکنش مالی (سالم و ناسالم) جهت پیش بینی تقلب در یک تراکنش با حل مساله طبقه بندی می باشد. الگوریتم های یادگیری ماشینی که باید بکار گرفته شوند عبارتند از: روش رگرسیون منطقی، روش بیز ساده و روش ماشین بردار پشتیبان، درخت تصمیم و یادگیری جمعی (Bagging، Boosting، Pasting).

قابلیت های برنامه

- ۱- خواندن اطلاعات دو فایل data1.csv و data2.csv از صفحه درس در سامانه lms و تشکیل یک DataFrame با استفاده کتابخانه pandas (ستون آخر نشان دهنده شماره کلاس نمونه است).
- ۲- رسم شماره کلاس نسبت به مقدار هر ویژگی در نمودارهای مختلف.
- ۳- رسم هیستوگرام برای ویژگی time برای کلاس های ۰ و ۱ (همانند شکل ۱).
- ۴- رسم هیستوگرام برای ویژگی Amount برای کلاس های ۰ و ۱.
- ۵- استانداردسازی ویژگی ها در صورت نیاز
- ۶- افزایش داده ها به ۳۰ درصد تست و ۷۰ درصد آموزشی.
- ۷- آموزش و آزمون طبقه بندی (رگرسیون منطقی، بیز ساده، ماشین بردار پشتیبان و درخت تصمیم) به صورت مجزا و رسم ماتریس درهم آمیختگی (confusionmatrix) متناظر با داده های آموزش و آزمون.
- ۸- رسم نمودار ROC برای هر اجرا.
- ۹- انجام عملیات فوق با اعمال الگوریتم PCA و انتخاب ۱۰ ویژگی مهمتر.
- ۱۰- استفاده از الگوریتم های یادگیری جمعی و مقایسه نتیجه آن با روش های فوق با رسم نمودار ROC و



شکل ۱: هیستوگرام time برای داده هایی با شماره کلاس ۱، bin=۱۰