

به نام خدا

## فاز شناخت و فهم مسئله پروژه درس مبانی داده کاوی 1401-02

اعضای گروه:

امیررضا حسینی – پاشا احمدی

عنوان پروژه:

بررسی رفتار کاربران در یک بانک جهت افتتاح حساب

<https://www.kaggle.com/datasets/janiobachmann/bank-marketing-dataset>

الف – مشکل چیست؟ (توضیح: چرا این مسئله یا داده تعریف شده است در واقع اگر این پروژه انجام نمی شد چه مشکلی، حل نشده باقی می ماند.)

امروزه، باتوجه به وجود طیف وسیع مشتریان و تنوع سلاقی و خواسته های گوناگون میان مؤسسات مالی و آنها، بیش از پیش نیاز به برنامه ریزی های کارآمد در مقیاس کلان حس می شود. از این رو، بانک های مختلف، با برنامه ها و پیشنهادهای تجاری زیاد پا به عرصه گذاشته تا گوی رقابت را از یکدیگر برابند. اما یکی از اساسی ترین چالش ها در این بازار تجاری، نبود سیستمی با راهبردی کارآمد در زمینه تبلیغات و کمپین های مختلف تبلیغاتی برای این بازار است. این سیستم باید بتواند بهترین و مفیدترین سناریوی ممکن را در اختیار بانک قرار داده تا بتواند با پیشنهاد دادن آن به هر یک از مشتریان، هم به آنها خدمت کرده و با راضی نگه داشتن آنان، سوددهی بلندمدت خود را به بالاترین حد ممکن برساند که باعث رضایت حداکثری هر دو طرف می شود. در این پروژه اصلی ترین هدف ما تلاش برای ساختن همچنین سیستمی می باشد.

ب- سوال داده کاوی و معیار ارزیابی آن چیست؟ (توضیح: قرار است چه مقدار (ستون یا اطلاعات) را پیش بینی (تخمین یا برآورد) کنید؟ و قصد دارید از چه اقلام اطلاعاتی برای این پیش بینی استفاده کنید؟ معیار ارزیابی چیست و ایده آن چقدر است؟)

به طور کلی، مجموعه داده های بازاریابی می توانند برای دو هدف تجاری مختلف استفاده شوند:

- پیش بینی نتایج یک کمپین بازاریابی برای هر مشتری و روشن شدن عواملی که بر نتایج کمپین تأثیر گذارند. این به کسب و کارها کمک می کند تا راه هایی برای بهبود کمپین های بازاریابی خود پیدا کنند.
- شناسایی گروه های مشتریان با استفاده از داده های مشتریانی که به وام ترم های سپرده علاقه مند شده اند. این به کسب و کارها کمک می کند تا پروفایل مشتریانی که احتمال خرید یک محصول خاص را دارند شناسایی کنند و کمپین های بازاریابی هدفمندتری را طراحی کنند.

در این پروژه ما قصد داریم تا با انواع تکنیک های داده کاوی مانند بصری سازی، آنالیز، شناسایی الگوها و روابط، یادگیری ماشین و ... به تحقق این اهداف دست یابیم تا بهترین و بهینه ترین کمک را به سوددهی بانک در استراتژی های کمپین های تبلیغاتی آتی کرده باشیم.

### ج- مشخصات دیتاست و ویژگی های موجود در آن

این مجموعه داده تشکیل شده از ۳ گروه داده می باشد:

داده های مربوط به مشتریان شامل:

- سن: (عددی)
- شغل: نوع شغل (دسته ای: 'ادمین', 'کارگر', 'کارآفرین', 'خانه دار', 'مدیر', 'بازنشسته', 'خود اشتغال', 'خدمات', 'دانشجو', 'تکنسین', 'بیکار', 'نامشخص')
- وضعیت تأهل: وضعیت تأهل (دسته ای: 'طلاق گرفته', 'متأهل', 'مجرد', 'نامشخص'),
- تحصیلات: (دسته ای: سیکل، دیپلم، عالی، و نامشخص)
- قرض الحسنه: آیا شخص حساب قرض الحسنه دارد؟ (دسته ای: 'خیر', 'بله', 'نامشخص')
- مسکن: آیا شخص وام مسکن دارد؟ (دسته ای: 'خیر', 'بله', 'نامشخص')
- وام شخصی: آیا شخص وام شخصی دارد؟ (دسته ای: 'خیر', 'بله', 'نامشخص')
- موجودی: کل موجودی حساب شخص. (یک متغیر عددی که می تواند برحسب بدهکاری یا طلبکاری منفی یا مثبت باشد).

داده های مربوط به آخرین تماس با هر مشتری در کمپین فعلی:

- نوع بستر ارتباط تماس (دسته ای: 'گوشی موبایل', 'تلفن ثابت')
- آخرین ماه تماس در سال (دسته ای: 'ژانویه', 'فوریه', 'مارس', ..., 'نوامبر', 'دسامبر')
- آخرین روز هفته تماس (دسته ای: 'دوشنبه', 'سه شنبه', ..., 'چهارشنبه', 'پنجشنبه', 'جمعه')
- مدت زمان آخرین تماس گرفته شده با مشتری بر حسب ثانیه (عددی).
- تعداد تماس های انجام شده در طول این کمپین و برای این مشتری (عددی، شامل آخرین تماس)
- تعداد روزهایی که پس از تماس آخرین با مشتری از کمپین قبلی گذشته است. (عددی؛ عدد 999 به معنای عدم تماس قبلی با مشتری است)
- تعداد تماس های انجام شده قبل از این کمپین و برای این مشتری (عددی)

- نتیجه کمپین بازاریابی قبلی (دسته‌ای: 'شکست', 'ناموجود', 'موفقیت')
- برچسب خروجی (Y) شامل:
- آیا مشتری سپرده کوتاه‌مدت باز کرده است؟ (باینری شامل بله و خیر)

#### د- فعالیت پیش رو و بیان مختصر ایده تیم

در این پروژه قرار است با استفاده از آنالیز داده‌های نامبرده و اعمال الگوریتم‌های داده‌کاوی مختلفی مانند شبکه‌های عصبی، درخت تصمیم و یا SVM روی آنها، الگوهای رفتاری کاربران را پیدا کرده و کشف الگوهای رفتاری بین آنها را بررسی کنیم. سپس با استفاده از رویکردهای دسته‌بندی و خوشه‌بندی و نتیجه‌گیری‌های مرتبط برای برنامه‌ریزی‌های کلان آتی بانکی می‌توانیم به بهینه‌ترین طرح‌ها و پیشنهادات برای طیف وسیع سلايق و خواسته‌های هر مشتری برسیم. در نتیجه با انجام این کار در این پروژه، می‌توانیم سطح رضایت کاربران و تصمیمات مدیریتی و برنامه‌ریزی را با توجه به نیازهای کاربران بهبود بخشیم تا در آینده نیز از استراتژی‌های کارآمدتری برای بازاریابی استفاده کنیم و همچنین سوددهی را برای مؤسسه بانکی بهینه کنیم.