

Analisis Pengguna Kartu Kredit di Bank

Link github saya : <https://github.com/Askar2002/Analisis-Pengguna-Kartu-Kredit-di-Bank.git>

Lembaga perbankan memenuhi berbagai kebutuhan keuangan, tetapi tidak semua produk yang ditawarkan ideal bagi semua pelanggan. Untuk melayani pelanggan mereka dengan baik, departemen pemasaran dan analitik di bank memiliki tugas mengidentifikasi kelompok yang lebih memungkinkan kebutuhan produk tertentu. Hal ini sangat penting di analisis terutama di bidang keuangan, karena produk memiliki potensi kerugian jika tidak diperlakukan secara bertanggung jawab oleh perusahaan dan konsumen.

Adapun permasalahannya :

1. Produk apa yang paling umum digunakan oleh nasabah bank kami?
2. Apakah nasabah dengan tenor bank yang lebih lama sering menggunakan kartu kredit?
3. Apakah segmen nilai berbeda dalam jumlah produk yang mereka gunakan?
4. Atribut pelanggan apa yang paling prediktif dalam penggunaan kartu kredit?

Berikut ini langkah-langkah untuk menjawab permasalahannya :

Business Understanding

Di lihat dari sudut pandang pelanggan, mengapa mereka menggunakan kartu kredit. Keuntungan dari kartu kredit ini adalah lapisan keamanan tambahan terhadap penipuan, waktu tambahan untuk membayar pembelian besar, dan kemudahan sengketa biaya dengan vendor. Alasan ini yang menjadi kartu kredit populer di finansial produk.

Data Understanding & Preparation

Dalam data understanding dan preparation saya mengeksplorasinya secara bersama karena keduanya terkait erat dan sering kali ada banyak pergantian diantara kedua langkah tersebut.

Data ini saya menggunakan dari <https://www.kaggle.com/c/santander-product-recommendation> . Ini mencakup indikator biner untuk setiap produk yang dapat dimiliki pelanggan dan atribut pelanggan seperti : Pendapatan rumah tangga bruto, segmen pelanggan, jenis kelamin, usia, dan vintage.

Bagian dari data preparation, dataset kami buat, khususnya untuk pemodelan dimana kita removed or imputed missing values, created a “number of product” variable, dan sampel dari total pengguna kartu kredit dan tidak pengguna kartu kredit.

Disamping dari understanding dataset, ada beberapa hal yang harus diketahui seperti :

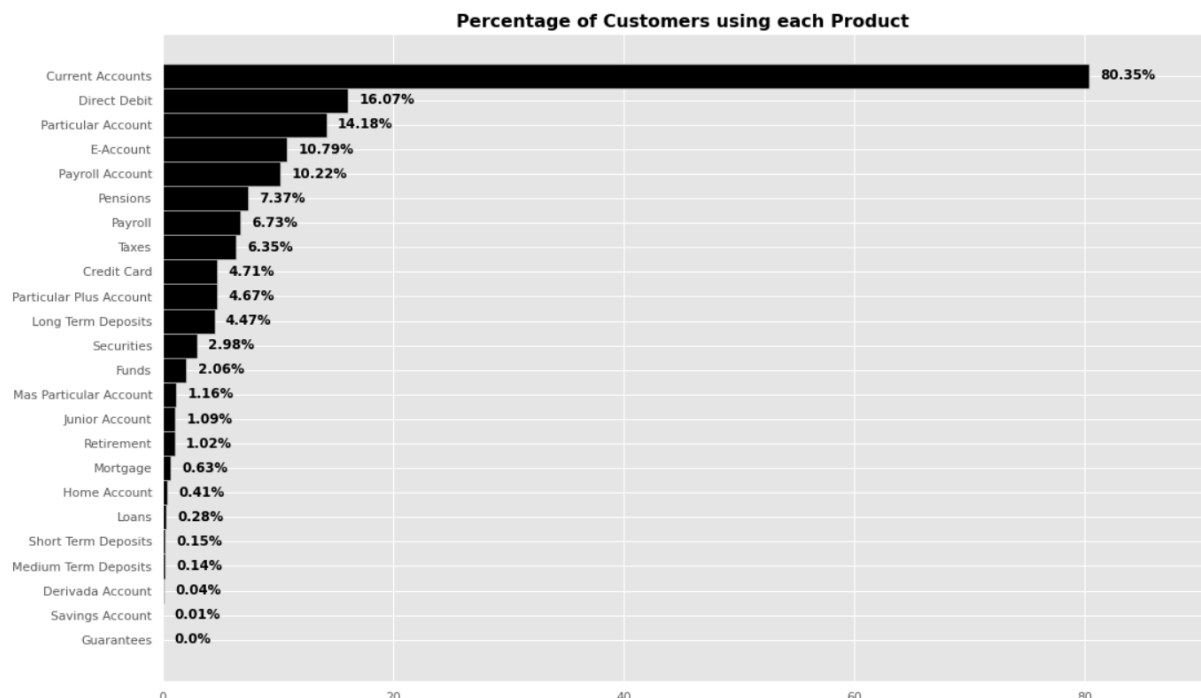
- Sistem apa yang menghasilkan data ?
- Apa yang sudah kita ketahui tentang variable dan pelanggan dalam data ?
- Berapa banyak pelanggan kami yang termasuk dalam data atau tidak termasuk?
- Mana variable yang mudah digunakan untuk tujuan pemasaran?

Data Modeling

Dalam pemodelan data ini saya menggunakan code dari github <https://github.com/LDMartinezSpecia/BancoSantander> . Disini saya menjawab pertanyaan permasalahan di awal :

1. Produk apa yang paling umum digunakan oleh nasabah bank kami?

Saya menjawab pertanyaan tersebut membuat suatu following plot :

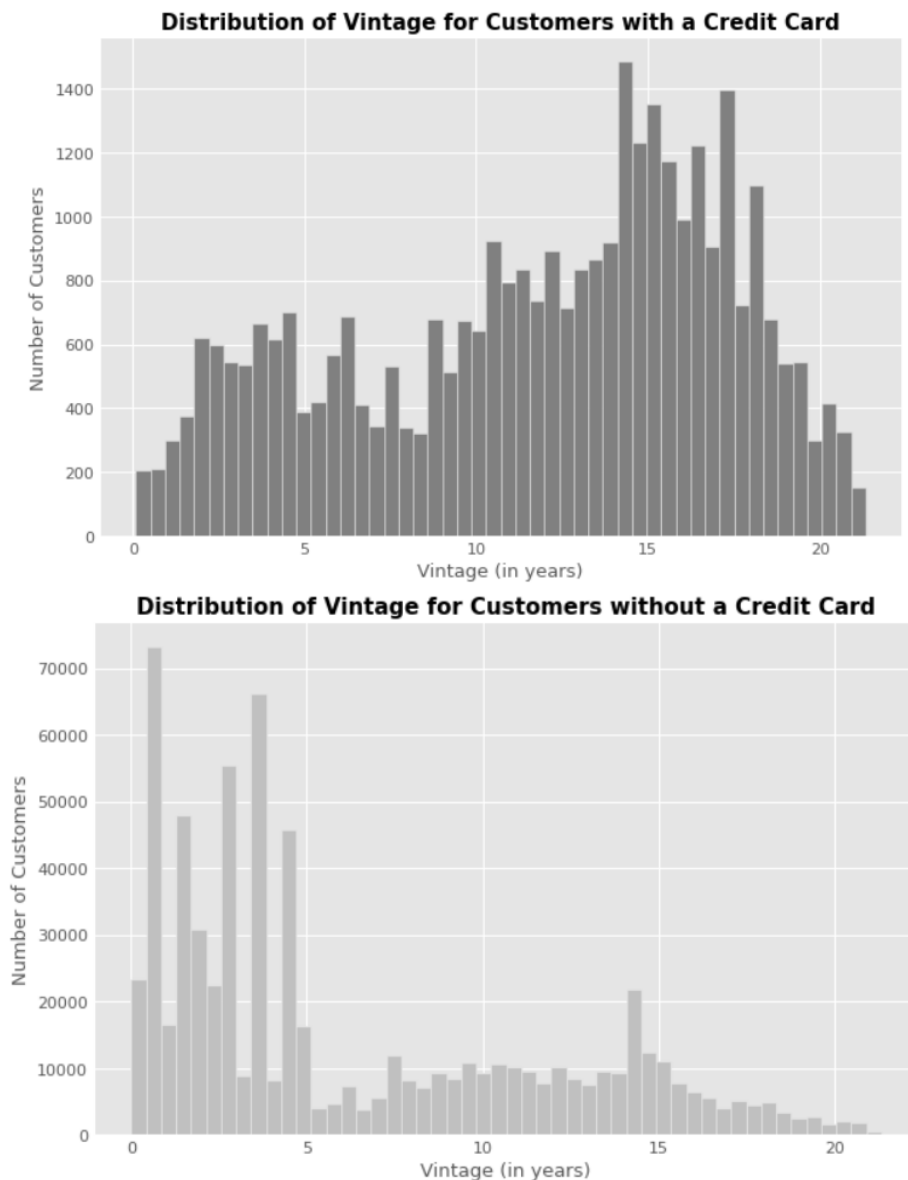


Dari plot diatas (menggunakan matplotlib library in python) bahwa “current account” sejauh ini produk yang paling umum digunakan, dimana lebih dari 80% pelanggan di dataset menggunakannya. Selain itu, penggunaan produk relative rendah dengan 4 produk selanjutnya sekitar 10% digunakan.

Berdasarkan plot ini initially clear mayoritas nasabah punya single produk (Summary statistic reveal the mean to be 1.7 products and the median to be 1). Sepintas tampak ada peluang memperdalam hubungan nasabah melalui penjualan silang produk kepelanggan internal, tetapi beberapa mungkin sulit dijual secara luas karena saat ini dipegang oleh sebagian kecil pelanggan.

2. Apakah nasabah dengan tenor bank yang lebih lama sering menggunakan kartu kredit?

Vintage didefinisikan sebagai “the number of month since a customer began their relationship with the bank” (I Changed to years rather than month in the visual below). Untuk menjawab permasalahan tersebut data dipisahkan oleh pengguna kartu kredit dan non kartu kredit dan distribusi vintage diplot :

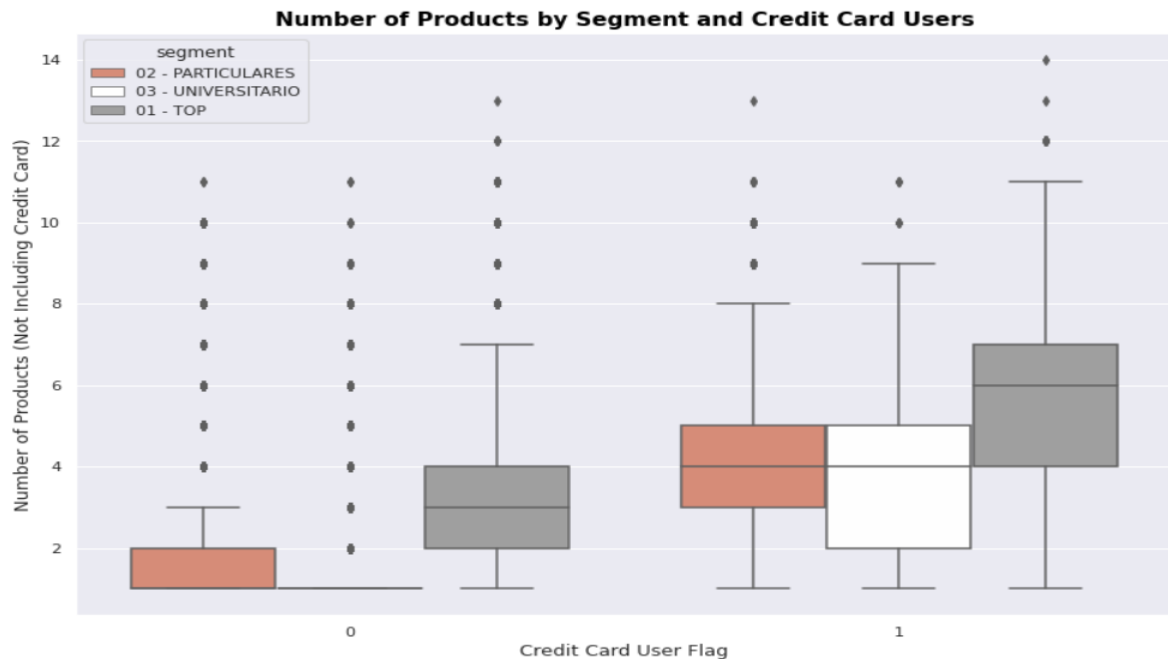


Perbandingan keduanya menunjukkan bahwa distribusi pelanggan dengan kartu kredit bergeser ke arah vintage yang lebih tinggi. Untuk lebih memperkuat poin ini, median untuk pelanggan dengan kartu kredit adalah 14,7 tahun sedangkan vintage median untuk rekan-rekan non kartu kredit mereka adalah 4,8. Pertanyaan wajar berikutnya mungkin apakah seorang nasabah telah berada di bank lebih lama karena mereka menggunakan kartu kredit atau apakah mereka menggunakan kartu kredit dengan bank karena mereka telah menjalin hubungan yang lama.

Meskipun kita tidak akan melihat terlalu dalam ke jawaban itu, kemungkinan ada efek kartu kredit menjadi produk yang cenderung digunakan oleh orang-orang yang lebih tua, mungkin juga ada efek bahwa pelanggan senang dengan keuangan mereka. lembaga lebih mungkin untuk tinggal lebih lama dengan lembaga itu dan karena itu menggunakan lebih banyak produk. Terlepas dari arah kausalitas, jelas bahwa nasabah yang memiliki kartu kredit rata-rata telah berada di bank lebih lama daripada yang tidak.

3. Apakah segmen nilai berbeda dalam jumlah produk yang mereka gunakan?

Bank dan bisnis lainnya akan sering mengklasifikasikan pelanggan mereka ke dalam segmen pelanggan serupa untuk tujuan membedakan layanan dan penawaran produk. Segmen tersebut tersedia dalam dataset ini, dan meskipun kami tidak mengetahui aturan yang digunakan untuk menentukan segmen ini, kami menyimpulkan dari namanya bahwa mereka dibagi menjadi "bernilai tinggi" (01-TOP), "reguler" (02-PARTICULARES), dan "mahasiswa" (03-UNIVERSITARIO). Plot berikut melihat distribusi jumlah produk (tidak termasuk kartu kredit) untuk masing-masing segmen, selanjutnya dibagi dengan apakah mereka memegang kartu kredit:



Sejauh ini, telah mempelajari beberapa hal menarik: pelanggan dengan kartu kredit cenderung memiliki vintage yang lebih lama (dan juga usia yang lebih tinggi), mereka juga lebih cenderung berada di segmen bernilai tinggi dan memiliki jumlah produk yang lebih banyak. Selain itu, kartu kredit tampaknya menjadi produk pembeda yang penting karena pelanggan yang memilikinya juga cenderung memiliki lebih banyak produk daripada rekan non-kartu kredit mereka di segmen yang sama. Hal ini terutama berlaku dalam segmen mahasiswa.

4. Atribut pelanggan apa yang paling prediktif dalam penggunaan kartu kredit?

Untuk ini saya memilih menggunakan run a decision tree dimana semua variable pelanggan. Memprediksi pengguna kartu kredit (class 1 seperti diagram dibawah). Setelah pemilihan iterasi dan pemilihan variable, pohon keputusan berikut dipilih untuk interpretasi lebih lanjut :

- Dia punya a higher number of product
- Aktif akunya
- Punya direct debit account
- Punya longer vintage

Secara khusus, jumlah produk yang dimiliki pelanggan adalah prediktor yang jauh lebih kuat daripada variabel lainnya (memiliki 3 atau lebih produk sebagai aturan pertama dan terkuat). Yang menarik adalah bahwa satu-satunya variabel "demografis" yang dipilih oleh model adalah variabel yang terkait dengan usia akun (lokasi, jenis kelamin, dan pendapatan yang memiliki pengaruh kecil).

Dari perspektif penjualan, gambaran yang mulai kita lihat adalah bahwa fokus atau upaya penjualan pada pelanggan yang memiliki banyak produk di bank dan telah bersama kami untuk jangka waktu yang lebih lama dapat meningkatkan peluang keberhasilan kampanye penjualan. Setelah memvisualisasikan variabel-variabel ini sebelum pemodelan, temuan dari pohon tidak akan mengejutkan kami.

Deployment

Meskipun "penyebaran" menyiratkan otomatisasi model dalam pipa untuk membuat keputusan, ini bukan satu-satunya cara model dapat digunakan untuk mengekstrak nilai dari data. Itu tentu saja merupakan kasus penggunaan yang canggih dan berharga, tetapi dalam istilah yang lebih sederhana, langkah penerapan juga dapat menjadi tempat kami mengambil temuan apa pun dan menerapkannya pada bisnis untuk mendapatkan manfaat nyata.

Misalnya, jika pertanyaan yang kami coba pecahkan datang langsung dari manajemen senior, maka "penerapan" model ini hanya memerlukan presentasi, mirip dengan proyek ini, di mana kami menampilkan temuan kami dan membuat rekomendasi untuk membantu pengambilan keputusan. Terlepas dari apa arti penerapan dalam kasus spesifik Anda, fokusnya harus pada bagaimana analisis yang telah dilakukan dapat berdampak pada perusahaan.

Refrensi

<https://www.kaggle.com/c/santander-product-recommendation>

<https://github.com/LDMartinezSpecia/BancoSantander>