ML-banking Grobkonzept

Einfache use-cases für Bestandskunden einer Bank

1. Weboberfläche
   1. Bank soll festlegen können, was sie auswerten will
      1. Erste Stufe (einfach): vorgegebene Auswahl; zB
         1. Kunden die x kauften, kaufen üblicherweise auch y
         2. Welche Vorhersagen lassen sich aus den Mustern treffen (warten auf GPT-Vorschläge)
         3. …
      2. Zweite Stufe (komplexer): durch freie Sprach-/Texteingabe (zB. „zeig mir 3 Möglichkeiten für neue Angebote auf, welche ich meinen Kunden im dritten Quartal anbieten kann“
   2. Kunde soll seine Datenquellen eingeben/uploaden können (.csv files)
   3. Website-Ausgaben sollen nicht nur in Text sein, aber auch Grafiken
2. Datenbasis (.csv); Datensatz mit 10.000 Einträgen.  
   für jedes fett gedruckte Element 1 Datei anlegen, und diese mit dem Kundenstamm verbinden (via key)
   1. **Kundenstamm** aus Österreich (Key – Kundennummer, Vorname, Nachname, Wohnadresse, Postleitzahl, Stadt, Land, Geburtsdatum, 3 Geschlechter: Mann, Frau, Divers – wenige Prozent; Bildungsniveau: kein Schulabschluss, Lehre, Matura, Studium)
   2. **Bank-Produkte** je Kunde. Manche Kunden haben nur 1 Produkt (zB Girokonto), andere haben mehrere Produkte (zusätzlich Sparbuch – je Kunde bis zu 3 Sparbücher möglich, Finanzierungen f. Immobilien oder Konsumkredite, Wertpapierdepot, Kreditkarte, Debitkarte, etc.). Teile jedem Kunden zumindest 1 Produkt zu, manche Kunden können bis zu 5 Produkte haben
   3. **Transaktionstyp**: Ordne jeder Transaktion (Eingang oder Ausgang) einen Kanal zu: eBanking (Web), mobile Banking (App), physische Bankbesuche, Cashtransaktionen – sowohl am POS im Geschäft, oder Cash-Abhebungen am Bankomat.  
      Ordne jede Transaktion eines Kunden einer Kategorie zu:
      1. Einkommen (Gehalt, Dividenden, Zinsen, Transferleistungen wie Kindergeld, Arbeitslosengeld, Sozialhilfe, Rente, Steuerrückerstattungen, Unterhaltszahlungen, Schenkungen/Erbschaften, sonstige)
      2. Ausgaben (Miete, Mietnebenkkosten, Internet, Versicherungen wie Haftpflichtversicherung, Unfallversicherung, Lebensversicherung, Pensionsversicherung, Auto, Telefon, Renovierung, Supermarkt-Einkauf, Kleidung, Schuhe, Frisör, Pflegeartikel, Haustiere, öffentliche Verkehrsmittel, Sportausgaben, Freizeit, Arzt, Medikamente, Apotheke, Physiotherapie, Psychotherapie, Brille, Restaurantbesuche, Kino, Hobbies, Reisen und Urlaub, Bücher, Weiterbildungen, Kurse, Kinder – Taschengeld, Schulgeld, Altersvorsorge, Kreditraten, Ausgaben für Weihnachten oder Ostern, Spenden an gemeinnützige Organisationen)
   4. **Werbeaktionen** der Bank: Simuliere diverse Marketing- und Werbeaktionen der Bank; zB Werbe- und Social-Media Kampagnen zu diversen Produkten; derartige Kampagnen sollten ca. alle 1 Monate stattfinden
   5. **Cookie-Daten**: die Bank hat in ihrer eBanking-App und mobile-App Cookies hinterlegt, sodass sie auch das User-Verhalten besser analysieren kann. 🡪 später ggf. einbinden; keine Prio jetzt
   6. Zeitraum der Datenbasis: erfasse einen Zeitraum von 3 Jahren, Darstellung auf Tagesebene; ein Kunde kann täglich mehrere Transaktionen haben, oder über Wochen gar keine. Durchschnittlich sollte ein Kunde 50 Transaktionen pro Monat haben.
   7. Bezüglich der Werte der Transaktionen nimm einen Durchschnitt der üblichen Ein- und Ausgaben eines österreichischen Menschen. Nimm aber nicht nur „Durchschnittsmenschen“, sondern alle Personengruppen (Kind – bis 14, Jugendlicher – bis 18, Erwachsener – bis 60, Pensionist – ab 61).
   8. Weitere Infos bzw. Promt zur Datenbasis: mach die Daten nicht zu homogen bzw. gleichverteilt, sondern nimm ein paar Ausreisser mit in die Datensätze; zB manche Kunden haben kaum Transaktionen, andere sehr viele; hohe bzw. geringe Zahlungen/Umsätze; zeitliche Unterschiede wie manche haben nur alle paar Monate Transaktionen; etc. 🡪 ein 80% homogenes Bild, und 20% Ausreisser