

Minisystem rekomendacji dla centrum obsługi klienta



Mały system rekomendacji, czyli nisko wiszący owoc prowadzący do mierzalnej przewagi konkurencyjnej

W kontaktach sprzedażowych B2B duży wpływ na późniejsze decyzje ma wrażenie pierwszych sekund rozmowy. Call center, które na starcie trafnie proponuje konfigurację pojazdu dostawczego, skraca ścieżkę decyzyjną klienta i zmniejsza jego niepewność. W niniejszym tekście przedstawiam koncepcyjny, tani i prosty system rekomendacji dla call center sprzedającego piekarni pojazdy dostawcze do 3,5 t.

Architektura systemu rekomendacji jest celowo prosta, a jednocześnie bardzo skuteczna operacyjnie. Klucz stanowi filtr finansowy zgodny z deklarowanym limitem miesięcznej raty oraz ranking popularności proponowanego pojazdu.

Celem opracowania jest nie tylko opis metody, ale i argumentacja, dlaczego „prosty rower” bywa w realnej organizacji bardziej sprawczy niż „samochód wyścigowy”, którego budowa przeciąga się miesiącami. System ma być gotowy do wdrożenia w krótkim czasie, mierzalny i łatwy do skalowania; jego przewagą jest przewidywalność, dyscyplina decyzyjna konsultantów i odporność na „szum” w wąskich segmentach klientów.

Gotowa do zastosowania aplikacja, opisana w tej publikacji, napisana w języku Python, znajduje się w repozytorium GitHub pod adresem: <https://github.com/ff-wm/reco-A1>

Klient ID:		649				
		Pokaż TOP-5				
	index	segment	model	trim	financing	monthly_rate_est
0	97	sedan	C1 Sedan	base	lease_36	103.33
1	84	compact	B3 Active	comfort	cash	0.00
2	6	city	A1 CityGo	comfort	lease_48	42.97
3	65	compact	B2 Hatch	base	lease_36	77.50
4	85	compact	B3 Active	comfort	lease_36	85.25



Kontekst biznesowy i motywacja

Wyobraźmy sobie call center firmy pośredniczącej w sprzedaży lub dzierżawie samochodów. W klasycznym scenariuszu konsultant odbiera telefon od właściciela piekarni i słyszy pytanie: „Jaki pojazd będzie dla mnie

właściwy?” Jeśli odpowiedź jest nietrafiona, np. pada propozycja bardzo drogiego samochodu bez finansowania w sytuacji, gdy klient oczekuje leasingu z określoną ratą, rozmowa de facto natychmiast się kończy, choć grzecznościowo może jeszcze chwilę potrwać. Pierwsze trafienie jest kluczowe, ponieważ obniża wysiłek

poznawczy klienta i sygnalizuje dopasowanie: „Ci ludzie rozumieją moje potrzeby”. Odwrotna sytuacja uruchamia mechanizm unikania: klient nie wraca, bo doświadczył dyskomfortu i ma poczucie straty czasu.

Organizacje często odkładają budowę własnego systemu rekomendacji, czekając na „idealny” moment lub tłumacząc sobie: „jesteśmy sprawni, po co zmieniać to co działa dobrze?”. Ostatecznie powstają mrzonki o przyszłym wielkim systemie, który zrewolucjonizuje sprzedaż. Jest to kosztowne złudzenie. Biznes potrzebuje wyników tu i teraz, a nie demonstracji nowych technologii czy zaawansowanej sztucznej inteligencji. Idea „niskowiszących owoców” polega na wdrożeniu rozwiązań, które – mimo prostoty – natychmiast przynoszą efekty, poprawiają kluczowe wskaźniki: konwersję, średni przychód na kontakt czy skracają czas obsługi transakcji. „Rower” nie jest tak szybki jak sportowy samochód, ale trzykrotnie przyspiesza marsz i robi to za ułamek ceny samochodu wyścigowego: to metafora wartości dodanej prostych rekomendacji w porównaniu z zaawansowanymi modelami, które wymagają wielkich nakładów czasu, kapitału i pracy. Z moich obserwacji wynika, że wdrożenie najprostszego systemu rekomendacji przynosi nawet 20% wzrost sprzedaży ilościowej. Kolejne modyfikacje oraz na bazie tego rozwiązania wdrożenie super rozwiązań opartych o sztuczną inteligencję powoduje wzrost o niespełna 3-5%. Tak więc mały, tani, robiony za ułamek ceny „wielkiego rozwiązania” system rekomendacji może okazać się *game changerem* efektywności operacyjnej.

Ramy problemu i założenia danych

Rozważamy sprzedaż pojazdów użytkowych dla piekarni. Interesuje nas segment do 3,5 t, ponieważ dla piekarni ten przedział jest szczególnie atrakcyjny. Łączy on dostępność kierowców z prawem jazdy kategorii B z funkcjonalnością wystarczającą do codziennego dowożenia pieczywa. Zakładamy, że call center ma dostęp do bazy klientów, którzy dzwonią w sprawie nowych samochodów. Baza zawiera informacje o maksymalnym budżecie miesięcznym, czyli limicie raty leasingowej, podstawowe dane klientów takie, jak np. skala produkcji, region działania. Klienci z historią zakupową mają w bazie informacje o segmentach, modelach samochodów, wariantach wyposażenia i formach finansowania. Call center ma również katalog pojazdów, w którym są takie informacje jak: segment, model, wersja wyposażenia (trim), forma finansowania oraz estymowana cena i wynikająca z niej szacowana rata kredytu lub leasingu.

Ważne jest rozdzielenie dwóch ról danych:

- po pierwsze – pełnią one funkcję ogranicznika decyzyjnego w postaci budżetu klienta, który potem przełoży się na twardy filtr dla klientów;
- po drugie – dostarczają sygnału preferencji w postaci częstości zakupów na przekroju klas finansowych. Pozwala to budować rekomendacje, które są zarówno

wykonalne (mieszczą się w limicie raty), jak i zgodne z nawykami zakupowymi podobnych klientów.

Mechanizm rekomendacji

Metoda składa się z dwóch sekwencyjnych kroków, ale obie części tworzą jedną, zwartą procedurę rekomendacyjną.

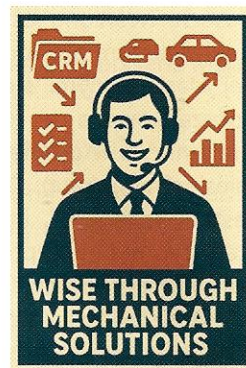
Najpierw dla konkretnego klienta wyznaczamy jego pozycję dochodową względem populacji. W praktyce najwygodniej jest podzielić rozkład deklarowanego limitu raty na sześć równolicznych przedziałów (*sextyle*). Innymi słowy dzielimy klientów na sześć klas ekonomicznych. Można wybrać dowolną ilość klas zamożności, liczba nie ma tutaj dużego znaczenia. Taki podział równoważy liczebności i ułatwia szacowanie stabilnych częstości zakupów. Następnie kandydatów do rekomendacji ograniczamy do tych konfiguracji pojazdu, których prognozowana miesięczna rata nie przekracza deklarowanego limitu klienta. Ten twardy filtr finansowy jest elementem krytycznym: usuwa przypadkowe i czasem kompromitujące propozycje call center, które „wysadzają w powietrze” pierwsze wrażenie i prowadzą do wycofania się klienta z transakcji.

W drugim kroku porządkujemy listę sześciu klas ekonomicznych, czyli zdolność budżetową klientów. Aby to zrobić, oceniamy każdy wariant pojazdu w kontekście klasy finansowej klienta poprzez jego popularność warunkową w każdej klasie, a konkretnie częstość zakupów w danym przedziale oraz popularność globalną, czyli częstość w całej populacji klientów. Sam sygnał informujący o najczęściej kupowanej konfiguracji w danym przedziale finansowym bywa niestabilny, gdy liczba obserwacji jest mała. Jeden „egzotyczny” zakup potrafi przesunąć ranking.

Dlaczego stosujemy wygładzanie addytywne, a nie na przykład średnią ważoną czy medianę?

Ponieważ dążymy do utrzymania stabilności pomimo rzadkości danych w niektórych klasach. Stałe wagi ignorują fakt, że klasy mają różną liczebność. M-estimate automatycznie skaluje wpływ globalu do wielkości próbki, dzięki czemu klasy z małą liczbą transakcji nie „przepalają” rankingu przypadkiem. Załóżmy, że klasa średnia zawiera tylko 5 transakcji, w obrębie tych transakcji jeden klient kupił 7 pojazdów indyjskiego koncernu Tata Motors. Jedna rzadka transakcja może zakłócić stabilność systemu rekomendacji w klasie zawierającej mało transakcji, dlatego konieczne jest stabilizowanie jej przez średnie globalne.

Celem systemu rekomendacji jest dostarczenie rankingu najbardziej popularnych pojazdów w określonych



klasach. Tworzona jest swoista lista top 5 samochodów z największym prawdopodobieństwem sprzedaży w danej klasie. Innymi słowy powstaje lista rekomendacji uporządkowana z rosnącym prawdopodobieństwem „trafienia”, a w drugiej kolejności – malejącym kosztem miesięcznym (rata). Dzięki temu konsultant otrzymuje zestaw propozycji, które są jednocześnie wiarygodne statystycznie i akceptowalne finansowo. Jak wspominałem, w przypadkach skrajnie rzadkich pojazdów na rynku, które pojawiły się w określonej mało licznej klasie stosujemy łagodzenie skrajności przez udział rankingu globalnego, co uniemożliwia dominację egzotycznych marek, będących swego rodzaju anomalią kulturową.

Intuicja i stabilność

Klienci reagują nie tylko na cenę i specyfikację, lecz na wrażenie dopasowania i oszczędności czasu. Pierwsza sensowna propozycja obniża niepewność i skraca dystans negocjacyjny. Z perspektywy statystycznej metoda wykorzystuje fakt, że rozkłady preferencji w obrębie zbliżonych ograniczeń finansowych są do siebie podobne. Domieszka popularności globalnej działa jak regulator, który „uspokaja” ranking, kiedy próbka lokalna jest zbyt mała. Z biznesowego punktu widzenia rozwiązanie jest transparentne: potrafimy wyjaśnić konsultantowi, dlaczego dany pojazd znalazł się wysoko. Kupowali go „tacy jak Pan/Pani”, jest więc popularny i spełnia założenia budżetowe, co sprzyja zaufaniu do call center i ułatwia proces. Każdy wie, że popularny model samochodu to łatwy dostęp do części zamiennych serwisu oraz duża liczba informacji od użytkowników w sieci.

Wdrożenie operacyjne bez kodu

Nie jest konieczne, by konsultant rozumiał szczegóły algorytmu. Z jego perspektywy interfejs to pole do wpisania identyfikatora klienta; wtedy pojawia się lista pięciu propozycji pojazdów z prognozowaną ratą i krótkim uzasadnieniem. Krytyczne z punktu widzenia psychologii sprzedaży jest udzielenie konkretnych rekomendacji. Rekomendacja otwiera rozmowę, ale nie zamyka jej – konsultant zadaje dwa, trzy pytania weryfikujące np. preferencje nadwozia czy dostępność kierowcy z prawem jazdy kat. B, po czym potwierdza wybór i proponuje wariant zamienny w tej samej klasie finansowej. Taki schemat ogranicza kompromitujące propozycje, „ofertowe strzały w płuć” i buduje wrażenie profesjonalizmu i pewności siebie.



Pomiar efektu w warunkach produkcyjnych

Organizacja powinna oceniać system w dwóch perspektywach.

Po pierwsze – przez pryzmat wskaźników operacyjnych: konwersja wpływu kontaktu na sprzedaż, przychód na kontakt, czas do decyzji na kontakt.

Po drugie – z punktu widzenia jakości rankingu: czy kupiony pojazd znalazł się w pierwszej piątce propozycji, czyli liście z rekomendacji i na jakiej pozycji na tej liście.

Najbardziej pragmatyczny eksperyment to prosty podział konsultantów call center na grupę testową i kontrolną. Należy zebrać dane „przed” i „po” wdrożeniu w tej samej testujących nowe rozwiązanie pracowników albo prowadzić równoległy test efektywności w tych samych oknach czasowych.

Przy normalnych rozkładach i porównaniu „przed/po” stosujemy test tStudenta dla prób zależnych; przy odchyleniach od normalności – test nieparametryczny MannaWhitneya lub permutacyjny. Kryterium decyzyjne przyjmujemy konserwatywnie (np. poziom istotności 0,05), a oprócz wartości p raportujemy także efekty wielkości (Cohen’s d), co ułatwia komunikację wartości biznesowej. Mówiąc prościej – wdrożenie prostego systemu rekomendacji również wymaga przetestowania na małej próbie pracowników call center.

Ryzyka, etyka i zarządzanie zmianą

Prostota nie zwalnia z odpowiedzialności. Wrażliwe jest zwłaszcza to, w jaki sposób kategoryzujemy klientów – należy ograniczyć się do korelatu finansowego transakcji (limit raty), a nie pośrednich cech społecznodemograficznych, które mogłyby wprowadzać nieuprawnione uprzedzenia. System nie może „karać” klienta za odmienną poza uzasadnieniem ekonomicznym. Innymi słowy, jeżeli obywatel Indii chce kupić samochody Tata nie warto przekonywać go, że „komputer mówi nie”. Równie istotne jest zapobieganie nadmiernej eksploracji niszowych konfiguracji. Zabezpiecza to mechanizm wygładzania popularnością globalną. Wreszcie, wdrożeniu powinna towarzyszyć krótka ścieżka szkoleniowa dla konsultantów – z naciskiem na zrozumienie zasad działania i świadome formułowanie uzasadnień dla klienta.

Co dalej: Modelowanie sekwencji zakupów

Kiedy prosty system rekomendacji zacznie przynosić wyniki, można stopniowo zwiększać jego złożoność. Naturalne kierunki to uwzględnienie kontekstu sezonowości, regionu czy wykorzystanie logistyki tras. Można pójść w kierunku modelowania sekwencji zakupu aut, czyli zależność kolejnych pojazdów w czasie. Modelowanie sekwencji zakupów to dodanie do rekomendacji informacji co zwykle wybierają klienci jako następny

pojazd po poprzednim wyborze i kiedy to zwykle robią. Innymi słowy: nie tylko to co jest popularne w określonej klasie budżetowej teraz, ale też co pewnym prawdopodobieństwem zwykle następuje po określonej wcześniejszej transakcji, np. po 36 mies. leasingu furgonu Atom firmy Ford klienci najczęściej przechodzą na wersję chłodniczą AtomFrost+. Przypis „co po czym” polega na prostej zasadzie częstości przejść, np. z segmentu A na B w tej samej klasie budżetowej z lekkim wygładzaniem, a ich wpływ rośnie im świeższy i dłuższy jest dziennik historii klienta. Metoda ta jest dość popularna wśród producentów samochodów, mówi się, że samochody rosną¹.

Dalszy rozwój: uczenie do rankingu Learning-to-Rank

Prosty, działający i zarabiający dodatkowe pieniądze model rekomendacji można rozbudowywać w drodze ewolucji. Kluczowe jest zachowanie rygoru mierzalności: żadna rozbudowa nie powinna osłabiać prostoty interpretacji w rozmowie z klientem.

Traktujemy rekomendacje jak problem ustawiania kolejności: model uczy się, które pojazdy powinny być wyżej dla danego klienta. Zamiast przewidywać „czy kupi”, przewidujemy relację „Ford przed Renault”. W praktyce karmimy model cechami klienta: budżet, cechami pojazdów z historii, finansowaniem, ratami i etykietami z logów, czyli informacje o przebiegu procesu zakupowego².

Przykład: dla klienta z budżetem raty miesięcznej 1800 PLN model uczy się, że „SUV-lease_48” zwykle wygrywa z „wagon-loan_60”, więc w TOP-5 przesunie SUV-y wyżej, bo historycznie częściej kończą się kontaktem lub zakupem.

Dalszy rozwój: mechanizm bandyty multi-armed bandit

To takie podejście rozwoju ewolucyjnego sprowadzające się do dwóch równoważących się kierunków:

- eksploatację – mówimy modelowi: pokazuj to, co działało dotąd
- eksplorację – mówimy modelowi: czasem pokaż coś mniej znanego, żeby sprawdzić czy nie działa lepiej.

Każdy ten kierunek to wariant rekomendacji, konkretna konfiguracja albo reguła układania listy. Proste algorytmy: ϵ -greedy, np. w 5% przypadków testujemy nowość, UCB promuje warianty z dużą niepewnością, ale potencjałem, Thompson Sampling próbuje proporcjonalnie do szacowanej skuteczności.

Na przykład: mamy 3 strategie układania listy:

- (S1) czysta popularność (czyli to co mamy opisane w tym artykule),
- (S2) popularność + sekwencje „co-po-czym”,
- (S3) popularność + większy nacisk na leasing_36, który akurat jest wyjątkowo zyskowny dla call center.

Bandit przydziela rozmowy konsultantów do S1/S2/S3; po każdej rozmowie aktualizujemy „wygraną”. Z czasem algorytm coraz częściej wybiera strategię, która daje najwyższy zysk, ale nadal okazjonalnie testuje pozostałe, by wyłapać zmianę trendu.

Reasumując: podejście *multi-armed bandit* to prosta metoda uczenia online, która rozdziela rozmowy między kilka strategii (np. S1/S2/S3), po każdej rozmowie aktualizuje ich „wygraną” (klik/umówienie/zakup), a z czasem coraz częściej wybiera tę strategię, która daje najwyższy zysk – jednocześnie czasem testując pozostałe, żeby nie przegapić lepszej opcji.

Wnioski

Nawet najprostszy opisany tutaj system rekomendacji – dopasowanie budżetowe i ranking popularności ze stabilizującym wygładzaniem – może istotnie poprawić wynik sprzedaży call center, podnosząc trafność pierwszej propozycji, a tym samym konwersję sprzedaży.

Przewaga tego systemu rekomendacji wynika z trzech cech:

- jest szybki do wdrożenia,
- transparentny w wyjaśnieniu,
- odporny na rzadkość danych w wąskich segmentach.

Wbrew intuicji, „zabawkowy” nie znaczy niepoważny: w wielu kontekstach to właśnie on jest pierwszym krokiem, który pozwala w ogóle zacząć podróż ku bardziej zaawansowanym formom personalizacji i optymalizacji przychodów.

Wojciech Moszczyński

Wojciech Moszczyński – absolwent Katedry Ekonometrii i Statystyki Uniwersytetu Mikołaja Kopernika w Toruniu, specjalista z zakresu ekonometrii, finansów, *data science* oraz rachunkowości zarządczej. Specjalizuje się w optymalizacji procesów produkcyjnych i logistycznych. Prowadzi badania w obszarze rozwoju i zastosowania sztucznej inteligencji. Od lat zaangażowany w popularyzację *machine learning* oraz *data science* w środowiskach biznesowych.

¹ Golf VII (od 2012 r.) przeszedł na modułową platformę MQB – tę samą rodzinę, na której zbudowano Passata B8 (od 2014 r.). W praktyce oznaczało to „podłogę” i kluczowe elementy podwozia wspólne/zbliżone, więc Golf wyraźnie „urósł” w stronę dawnych rozmiarów Passata.

² Przykłady modeli: LambdaMART, XGBoost ranker, pairwise Logistic Regression. Metryki: NDCG@k, MAP, Hit@k