

Opis scenariusza

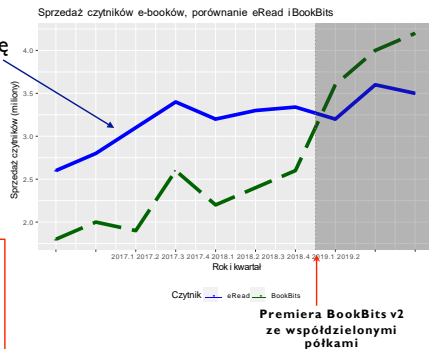
- WVCorp — firma zatrudniająca Ciebie (analityka danych).
 - Firma WVCorp dysponuje forami tematycznymi i dyskusyjnymi dla każdego produktu, gdzie klienci mogą dyskutować o problemach i cechach.
 - “Szum medialny” — wtedy, gdy jakiś temat na forum użytkowników ma bardzo duży poziom aktywności, co wskazuje na zainteresowanie użytkowników tym zagadnieniem.
- eRead — czytnik e-booków wyprodukowany przez firmę WVCorp.
- TimeWrangler — aplikacja do zarządzania czasem wyprodukowana przez firmę WVCorp.
- BookBits — czytnik e-booków wyprodukowany przez konkurencję.
- GCal — stworzona przez inną firmę rozproszona usługa kalendarzowa, którą można zintegrować z aplikacją TimeWrangler.

Przewidywanie szumu medialnego

Hipotetyczna prezentacja dla sponsora projektu przygotowana przez zespół analityków danych WVCorp

Stracona okazja

eRead — wcześniej najlepiej sprzedający się czytnik e-booków.



Szacunkowo utracono 25 mln dolarów przychodów ze sprzedaży produktu.

Określamy tu problem biznesowy stanowiący motywację dla projektu: firma WVCorp nie jest świadoma potrzeb klientów.

Czytnik eRead jest produktem firmy WVCorp (czytnikiem e-booków). Główną konkurencją w tym segmencie rynku jest czytnik BookBits.

Czy mogliśmy to przewidzieć?

- Dyskusje na forum produktu eRead:
 - Współdzielenie listy ulubionych książek ze znajomymi.
 - Udostępnienie książki grupie znajomych (kto pierwszy, ten lepszy).
- Zawsze, gdy pojawiały się pytania na ten temat, dyskusja ożywała.
 - Sugestie, obejścia, prowizorki, duża zgodność wśród klientów.
 - Wspólna półka (jak w BookBits) spełniały te wciąż powtarzane prośby.
- Wokół tego tematu narastał **szum medialny**! Ale zignorowaliśmy go. Albo przeoczyliśmy.
 - Ciągłe obserwowanie aktywności na forum jest bardzo pracochłonne.

Ustalamy tutaj, że problem można ewentualnie rozwiązać, jeżeli będziemy wykorzystywać informacje dostępne na forach produktu. Jednak w obecnym cyklu pracy monitorowanie forów stanowi problem. W prawdziwej prezentacji mógłbyś tu wstawić zrzut ekranu z odpowiedniego tematu na forum.

Cel: wczesne wykrywanie

- Przewidywanie, które tematy na forach produktu będą cechować się ciągłym szumem medialnym.
- Cechy wyczekiwane przez klientów.
- Istniejące cechy, które sprawiają problemy.
- Trwały szum: rzeczywiste, trwające potrzeby klientów.
- Problemy ani chwilowe, ani wynikające z trendów.

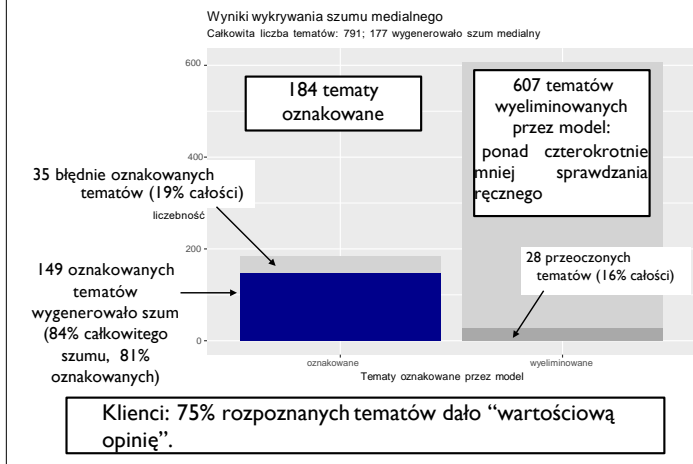
Wyjaśniamy tu cel projektu w kontekście zaistniałej potrzeby biznesowej. W rzeczywistej prezentacji warto byłoby wstawić kolejny zrzut ekranu z dyskusji na forum.

Badanie pilotażowe

- Zgromadzono dane na forum z trzech tygodni.
- W pierwszym tygodniu wytrenowano model, aby sprawdzić, które tematy będą szumieć medialnie w dwóch kolejnych tygodniach.
 - **Szum = utrzymywany wzrost 500+ aktywnych dyskusji na dany temat w ciągu dnia w porównaniu z pierwszym dniem w pierwszym tygodniu.**
- Przewidywane rezultaty porównano z rzeczywistością popularnymi wątkami.
- Opinie od pięcioosobowego zespołu menedżerów produktu — jak przydatne były wyniki?

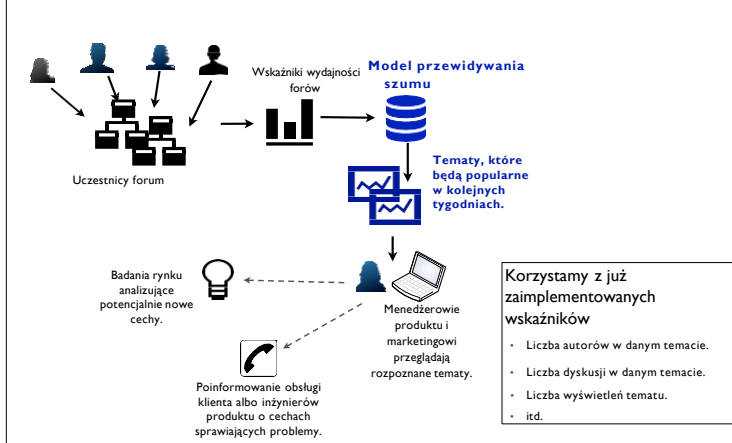
Szczegółowy, nietechniczny opis projektu.

Obiecujące wyniki



Określamy wynik projektu i wyjaśniamy, w jaki sposób i dlaczego okazał się udany. Zachowaj zwięzłość i unikaj szczegółów technicznych. Nie mów o precyzji/pełności, lecz o tym, w jakim stopniu zostało zredukowane obciążenie robocze menedżerów produktu/marketingowych, jak wiele potencjalnie przydatnych rezultatów model znalazł, a także jak wiele przeoczył.

Mechanizm działania



Szczegółowe, niezbyt techniczne wyjaśnienie użytych przez nas metod oraz ich miejsca w ogólnym procesie.

Model Buzz

- Model lasu losowego.
 - Głosuje wielu „ekspertów”.
 - Działa wydajnie na dużych zestawach danych.
 - Obsługuje dużą liczbę zmiennych wejściowych.
 - Kilka odgórnych założeń dotyczących interakcji zmiennych lub ich istotności.
- Bardzo dokładny model.

Zazwyczaj ta grupa odbiorców nie jest specjalnie zainteresowana wykorzystaną metodologią modelowania. Jeden szczegółowy slajd powinien wystarczyć; wyjaśnij tu, dlaczego wybrałeś ten model (dlaczego jest to „właściwy” algorytm). W niektórych sytuacjach wystarczy wspomnieć jedynie o najważniejszych kwestiach albo wręcz całkiem pominąć ten slajd.

Przykład: wczesne wykrywanie problemu

- Temat: Integracja TimeWrangler → GCal.
- Liczba dyskusji wzrosła od czasu wydania Gcal v7.
 - Zdarzenia GCal nie zawsze widoczne; niewłaściwy mechanizm ich oznaczania.
 - Zadania TimeWrangler wstawiane do złego obiektu GCalendar.
- **„Gorąco” na forach, jeszcze zanim zrobiło się „gorąco” w dziennikach zdarzeń obsługi klienta.**
 - Aktywność na forum pobudziła model dwa dni po aktualizacji GCal.
 - Obsługa klienta spóźniona o tydzień.

Dajemy przykład interesującego odkrycia w czasie realizacji projektu: problem, który został zidentyfikowany szybciej dzięki forum niż dziennikom zdarzeń pomocy technicznej. TimeWrangler to kolejny produkt firmy WVCorp, służący do zarządzania czasem; z kolei GCal jest popularną rozproszoną usługą kalendarzową stworzoną przez inną firmę. W prawdziwej prezentacji warto przygotować więcej przykładów.

Co dalej?

- Dalsze zmniejszenie obciążenia menedżerów produktu, przekazywanie im lepszych informacji od klientów.
- Nowe wskaźniki poprawiające przewidywania.
 - Rejestrowanie wzrostu/spadku aktywności dyskusji (a także jego szybkości).
 - Dlaczego rejestrują się nowi użytkownicy na forum? Jakiego zadają pytania?
- Cel: Wykrycie 98% nieuchronnego szumu przy 10% zdarzeń fałszywie pozytywnych.
- Skuteczne przekazywanie informacji o szumie do właściwych menedżerów produktu, marketingowych i grup obsługi klienta.

Wyjaśnij, w jaki sposób można usprawnić wyniki projektu, zaproponuj dalsze kroki.

Dziękuję