

Detektor spamu w wiadomościach SMS

Opis problemu:

Spam SMS, czyli niechciane wiadomości tekstowe są obecnie wysyłane masowo w celach reklamowych, bądź też oszukańczych. Wysyłanie tego typu powiadomień zwykle bywa bardzo uciążliwe dla odbiorców i podnosi ryzyko stania się ofiarą groźnych oszustw. Problem dotyczy wielu branż m.in. branży telekomunikacyjnej, czy finansowej. Wykrywanie spamu SMS ma na celu ochronę użytkowników przed niechcianymi wiadomościami, ryzykiem zostania oszukanym oraz poprawę jakości usług komunikacyjnych.

Cel projektu

Celem projektu jest opracowanie i wdrożenie skutecznego modelu klasyfikacyjnego, który będzie w stanie automatycznie identyfikować wiadomości SMS o charakterze spamowym.

Kto i w jaki sposób może skorzystać z tego modelu?

Model ten może zostać przeznaczony do użytku przez użytkowników telefonów komórkowych w celu ochrony przed niechcianymi wiadomościami, operatorów telekomunikacyjnych w celu zmniejszenia obciążenia sieci i poprawy jakości świadczonych usług, firmy e-commerce w celu lepszej komunikacji z klientami, eliminacji ryzyka wprowadzenia klientów w błąd przez oszustów.

Źródło danych

Podstawowy zbiór danych zawiera:

- Treść wiadomości SMS: tekst wiadomości
- Typ: klasyfikacja wiadomości jako spam lub ham (wiadomość bezpieczna, normalna)

W celu poprawy jakości modelu, zbiór danych po wcześniejszej analizie zostanie wzbogacony o dodatkowe kolumny wygenerowane na podstawie treści wiadomości. Kolumny te pomogą w wychwyceniu specyficznych cech charakterystycznych dla spamu.

Dataset zawiera 5169 rekordów i po jego refaktoryzacji (wzbogaceniu o dodatkowe kolumny) zostanie podzielony na **zbiór treningowy** o wielkości **3618** rekordów oraz **zbiór testowy** o wielkości **1551** rekordów

Link do datasetu: <https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>

Fragment datasetu:

ham spam	87% 13%	5169 unique values
ham		Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got a...
ham		Ok lar... Joking wif u oni...
spam		Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entr...
ham		U dun say so early hor... U c already then say...
ham		Nah I don't think he goes to usf, he lives around here though
spam		FreeMsg Hey there darling it's been 3 week's now and no word back! I'd like some fun you up for

Struktura pracy nad modelem

1. Pozyskanie i refaktoryzacja danych: Zebranie danych i przetworzenie ich na strukturę zgodną z wymaganiami.
2. Przetwarzanie danych: Usuwanie brakujących wartości, utworzenie dodatkowych kolumn, przygotowanie cech numerycznych, utworzenie wykresów i analiz.
3. Podział zbioru danych: Podział na dane treningowe (70%) i testowe (30%) w celu przeprowadzenia walidacji i kształcenia modelu.
4. Trenowanie modelu: Wybór algorytmu klasyfikacji, dostosowanie parametrów modelu do osiągnięcia wyznaczonych celów.
5. Walidacja i testowanie: Weryfikacja skuteczności na zbiorze testowym, obliczenie kluczowych miar efektywności i zależności modelu.
6. Wdrożenie: Konteneryzacja modelu do wykorzystania produkcyjnego.
7. Publikacja modelu i jego monitorowanie: Monitorowanie działania modelu na środowisku produkcyjnym i aktualizacja parametrów jeśli zajdzie taka potrzeba.