3. Projekt oraz implementacja modelu predykcyjnego wykrywającego transakcje noszące znamiona oszustwa

3.1. Czym jest detekcja oszustw transakcyjnych w kontekście kart kredytowych (Fraud Detection)

3.2. Charakterystyka użytego zestawu danych (dataset) – eksploracyjna analiza danych

3.3. Eksperyment pomocniczy - model optymalizacyjny wyłaniający najlepiej rokujących kandydatów spośród klasyfikatorów binarnych nadzorowanego uczenia maszynowego oraz ich złożeń

3.4. Eksperyment główny - budowa docelowego modelu predykcyjnego...............................

3.5. Interpretacja uzyskanych wyników

3.6. Podsumowanie zagadnienia

## 3. Projekt oraz implementacja modelu predykcyjnego wykrywającego transakcje noszące znamiona oszustwa

## 3.1 . Czym jest detekcja oszustw transakcyjnych w kontekście kart kredytowych (Fraud Detection)

Rosnąca liczba transakcji przeprowadzanych z wykorzystaniem kard kredytowych jako formy płatniczej w XXI wieku stała się miejscem, w którym wraz z ich wykładniczym wzrostem pojawia się proporcjonalna ilość prób wyłudzenia środków. Tworzy to swego rodzaju „raj” dla wszelkiej maści przestępców mających na celu wykorzystywanie luk w systemach płatniczych. Wydawcy kart kredytowych, właściciele biznesów honorujących taką formę płatności czy właściciele poszczególnych bankomatów lub nawet osoby fizyczne coraz częsciej padają ofiarami oszustów.

Szacuje się, że do 2028 roku dostawcy usług finansowych poniosą straty sięgające rzędu 40 miliardów dolarów, co jest znaczącym wzrostem wobec około 27 miliardów w 2018 roku [Nilson Report, 18.11.2019]. W samych Stanach Zjednoczonych rokroczny wzrost liczby oszustw z wykorzystaniem kart kredytowych prognozowany jest na poziomie 10%. Jednym z istotnych powodów takiej sytuacji jest stale poszerzający się rynek handlu informacjami wynikający z powszechności internetu, który przyczynia się do częstej kradzieży tożsamości skutkującej znaczącym uproszczeniem w konstruowaniu systemów pomagających w kreacji oraz automatyzacji oszustw.

Współmiernie do tej sytuacji rośnie zainteresowanie instytucji finansowych związane z etyką i moralnymi zasadami szeroko pojętej bankowości oraz budowaniem efektywnych systemów detekcji niepożądanych zachowań noszących znamiona oszustwa. Największe korporacje zaangażowane w tę gałąź finansów – American Express, Mastecard, Visa, Revolut czy UnionPay starają się skutecznie walczyć z problemami jakie niesie za sobą dynamiczna popularyzacja bezgotówkowych transakcji. Nadużycia będące nieodłącznym elementem stale rosnącej popularności kart kredytowych stanowią aktualnie jeden z filarów bezpieczeństwa nad usprawnieniem którego nieustannie pracują wyżej wymienione firmy. W związku ze stale rosnącą popularnością technologii związanych z uczeniem maszynowym, powszechnym dostępem do róznorodnych danych, analityką biznesową oraz inżynierią danych większość instytucji świadczących usługi związane z kartami kredytwymi zmienia swoje podejście i przykłada dużą wagę do optymalizacji bezpieczeństwa transakcji dokonywanych przy pomocy udostępnianych kart. Ten proces obejmuje zarówno klientów biznesowych (sklepy, punkty usługowe, centra wymiany walut) jak i indywidualnych.

Czym więc konkretnie jest oszutwo finansowe? (eng. Financial Fraud)

Oszustwo finansowe można scharakteryzować jako zespół procesów składających się na świadome działanie mające na celu uzyskanie korzyści finansowych w sposób nieuczciwy oraz niezgodny z obowiązującym prawem [Wójcik jw. (2008). *Oszustwa finansowe. Zagadnienia kryminologiczne i kryminalistyczne*, Warszawa]. W świecie finansów istnieje wiele różnych rodzajów nadużyć z których wyszczególniono najważniejsze pod kątem ich powszechości:

* Oszustwa podatkowe
* Oszustwa z wykorzystaniem kart kredytowych
* Oszustwa związane z przelewami (eng. Wire Fraud)
* Oszustwa związane z aktywami finansowymi (eng. Securities Fraud)
* Oszustwa inwestycyjne
* Oszustwa związane z ogłaszaniem upadłości finansowej

Oszustwo finansowe najczęściej jest odzwierciedlone w postaci fałszywej reprezentacji istniejących faktów określających działalność biznesową poprzez celowe zatajenie istotnych informacji lub dostarczenie fałszywych danych kontrahentowi. Ma to na celu uzyskanie bezprawnej korzyści finansowej, która potencjalnie nie zostałaby osiągnięta z wykorzystaniem legalnych mechanizmów. Najczęściej obserwowanym typem oszustw w minionej dekadzie jest to z wykorzystaniem kart kredytowych – poprzez posługiwanie się fałszywymi kartami lub poprzez kradzież danych prawowitych posiadaczy tych środków płatniczych i ich późniejsze wykorzystanie. Na tym właśnie przykładzie skupia się model będący częścią badawczą tej pracy.

Czym jest detekcja oszustw i na jakich mechanizmach wspomagających ją polegają korporacje?

Detekcja oszustw opisanych we wcześniejszym akapicie polega w dużej mierze na agregowaniu ogromnych ilości danych powiązanych z transakcjami oraz ich późniejszej integracji i dogłębnej, wnikliwej analizie. W obecnych czasach odbywa się to zazwyczaj z wykorzystaniem różnorakich metod uczenia maszynowego i technologii Big Data – w znacznym stopniu usprawnia to ten proces i zwiększa jego efektywność. To z kolei wpływa na wzrost bezpieczeństwa użytkowników takich środków płatniczych chroniąc ich pieniądze przed wyłudzeniami. Intensywny rozkwit możliwości technologiczych w drugiej dekadzie XXI wieku spowodował, że korporacje zaczęły w coraz większym stopniu polegać na automatyzacji procesów analizy danych płatniczych co przekłada się bezpośredno na możliwości wykrywania potencjalnych przestępstw z wykorzystaniem kart kredytowych.

Konkretne metody analizy danych oraz implementacji algorytmów uczenia maszynowego są jednak pilnie strzeżoną tajemnicą tych instytucji, gdyż jest to aspekt dający im potencjalnie konkurencyjną przewagę na rynku. Udostępnianie takich informacji mogłoby się przyczynić to spadku efektywności wdrażanych przez nie praktyk.

W dalszych podrozdziałach został opisany przeprowadzony w ramach pracy badawczej eksperyment, którego celem jest identyfikacja takich metod nadzorowanego uczenia maszynowego (lub ich złożeń), które na udostępnionym zestawie danych są w stanie osiągać jak najlepszą skuteczność detekcji oszustw przeprowadzanych z użyciem kart kredytowych.

**3.2 Charakterystyka wykorzystanego zestawu danych (dataset)**

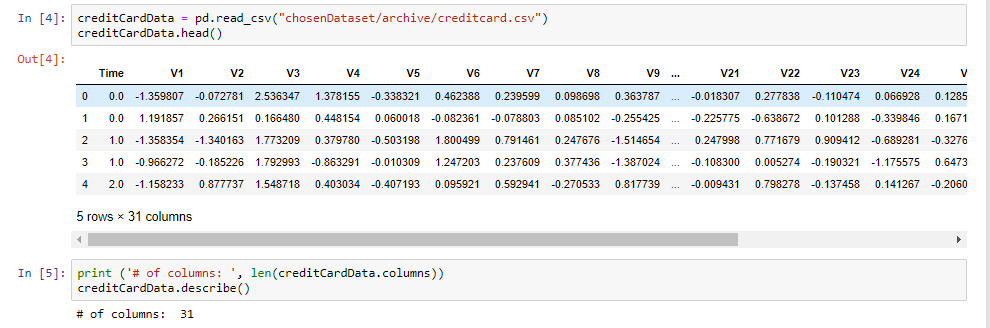
Eksperyment, którego celem jest estymacja wartości parametru opisującego czy dana transakcja jest fałszywa (eng. Fraudulent) czy zgodna z prawem został przeprowadzony na odpowiednio dopasowanym zbiorze danych. Zbiór ten na potrzeby pracy pozyskano z archiwum Université Libre de Bruxelles (ULB) [<https://mlg.ulb.ac.be>], a odpowiedzialny z jego stworzenie jest dział badawczy departamentu Computer Science. Zbiór danych jest dostępny pod adresem [https://www.**kaggle**.com/**mlg**-**ulb**/creditcardfraud] pod postacią pliku .csv.

MLG (Machine Learning Group) założone w 2004 roku przez profesora Gianluca Bontempi oraz profesora Toma Lenaertsa zajmuje się głównie przeprowadzaniem badań nad wykorzystaniem algorytmów uczenia maszynowego oraz inteligencji behawioralnej. Departament skupia się na analizie zbiorów, technologiach wynikania – casual inference, network inference, modelach wspomagania decyzji biznesowych oraz na anlizie behawioralnej znajdującejej szerokie zastosowanie między innymi w finansach, medycynie, cyberbezpieczeństwie czy badaniu dynamiki interakcji społecznych. Determinantem istotności departamentu dla wybranej dziedziny naukowej jest prężnie rozwijająca się liczba ukończonych projektów jak i mnogość zdobytych przez ostatnie lata nagród – „*Honorable Mention Research Paper at 2017 IEEE International Conference on Data Science and Advanced Analytics  (DSAA ’17)”, „Best paper award at the International Conference on Computational Collective Intelligence (ICCCI) 2020 conference.”* [https://mlg.ulb.ac.be/wordpress/awards/].

Wybrany zbiór danych charakteryzuje się 30 numerycznymi cechami których większość została przetransformowana przy pomocy techniki PCA (Prinicipal Component Analysis). Technika ta obfuskuje pozyskane dane w taki sposób, że przeprowadza na nich ortogonalną transformację liniową która przy zachowaniu referencyjności tych danych wobec siebie umieszcza je w nowej macierzy wartości. Jest to spowodowane wrażliwą naturą tych danych. W związku z zastosowaniem tej techniki otrzymujemy cechy opisujące o charakterze wartości zmiennych numerycznych [V1 => V28]. Dodatkowo, w zestawie danych mamy dwie pozostałe cechy –ciągłą – „Time”(Czas) oraz dyskretną „Amount” (Ilość). Cecha „Time” przedstawia wartości liczone w milisekundach, które reprezentują czas który upłynął pomiędzy transakcją opisaną danym wierszem a pierwszą transakcją obecną w zestawie danych. Cecha „Amount” reprezentuje natomiast wartość ilości (w walucie EURO), na którą została wykonana dana transakcja. Każdy wiersz obecny w datasecie reprezentuje indywidualną transakcję wykonaną za pomocą karty kredytowej. Ostatnią z cech które są dostępne w zestawie jest zmienna kategoryczna responsywna zdefiniowana za pomocą kolumny „Class”. Jest to zmienna będąca parametrem decyzyjnym w przeprowadzonej analizie. Przyjmuje ona wartości [0/1], gdzie „0” reprezentuje transakcję prawidłową natomiast „1” transakcję posiadającą znamiona oszustwa kredytowego.

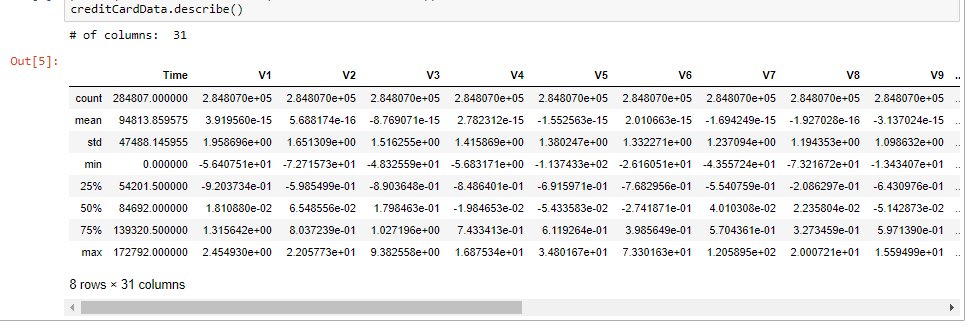
Dataset zawiera transakcje poczynione z wykorzystaniem kart kredytowych przez europejskich posiadaczy we wrześniu 2013 roku. Spektrum czasowe na które opiewają dane to 2 dni.

W celu zgłębienia informacji o wykorzystywanym na potrzeby eksperymentu zestawie danych przeprowadzono eksploracyjną analizę danych (EDA).



Rysunek 1: Liczba kolumn w zestawie danych oraz ich przykładowe wartości na podstawie kilku wierszy

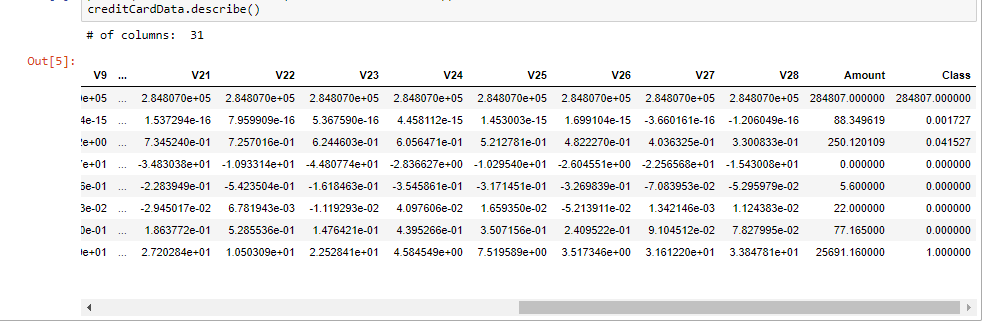
Na powyższym zrzucie ekranu przedstawiającym fragment kodu zaobserwowano jak kształtują się przykładowe wartości zmiennych obecnych w opracowanym zestawie danych.



Rysunek 2: Podstawowe statystyki opisowe zestawu danych

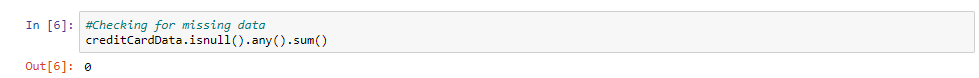
Na powyższym zrzucie ekranu przedstawiono wartości podstawowych statystyk opisowych. Dzięki nim uzyskano ogólny obraz dystrybucji danych na przestrzeni datasetu. Liczebność zbioru wynosi 284807 rekordów.

Po wstępnym przeprocesowaniu danych dokonano analizy zestawu pod kątem struktury danych powiązanych z transakcjami zawierającymi znamiona oszustwa kredytowego.



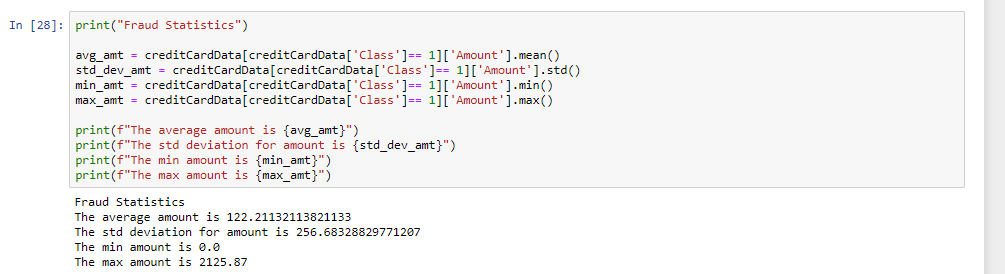
Rysunek 3: Kolumny Amount i Class

Z powyższego zrzutu wywnioskowano, że średnia wartość transakcji w zestawie wynosi 88.34 EUR(zaokrąglono). Minimalna wartość transakcji wynosi 22 EUR, a maksymalna 77.165 EUR.



Rysunek 4: Ilość danych nieistotnych(zerowych)

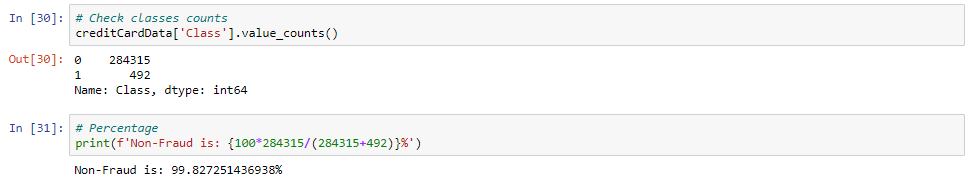
Sprawdzona została liczebność brakujących danych w zbiorze za pomocą funkcji .isnull(). W wyniku jej implementacji stwierdzono, że zbiór nie zawiera brakujących wartości w kontekście każdej kolumny oraz każdego wiersza.



Rysunek 5: Statystyki opisowe transakcji nieuczciwych

Poprzez analizę zmiennej decyzyjnej „Class” tam, gdzie jej wartość wynosi „1” zostały wyekstrahowane interesujące statystyki. Średnia wartość transakcji nieczuciwej wynosi ~122.22 EUR. Odchylenie standardowe tych wartości to ~256.7. Minimalna wartość w EUR takiej transakcji wynosi 0, natomiast maksymalna ~2125.9.

Kolejnym bardzo ważnym krokiem analizy jest wstępna ocena dystrybucji transakcji nielegalnych skonfrontowana z ilością transakcji prawidłowych. Dzięki dokonaniu takiej analizy uzyskano interesujące wartości.



Rysunek 6: Analiza dystrybucji transakcji nieuczciwych w stosunku do transakcji uczciwych na przestrzeni analizowanego zestawu danych

## Z powyższego zrzutu ekranu wyciągnięto wniosek, że dane w zbiorze są wysoce niezbalansowane. Oznacza to, że stosunek transakcji nieuczciwych („Class” = 1) stanowi jedynie promil w kontekście zawartości zestawu danych. Ilość transakcji potencjalnie nieszkodliwych wyznaczono na 99.82%. Taka sytuacja może potencjalnie skutkować stronniczym charakterem modelu, dlatego istotne jest wzięcie jej pod uwagę na jak najwcześniejszym etapie jego budowy, aby odpowiednio dobrać próbki uczące oraz testowe – by zawierały one pełne spektrum danych odzwierciedlające strukturę wartości w zestawie danych. Po wstępnej analizie otrzymanych rezultatów działania związane z eksploracyjną analizą danych zostały skierowane w stronę ich wizulizacji. Do tego celu wykorzystano biblioteki dające możliwość tworzenia Heat Map, wykresów punktowych oraz wykresów obrazujących jądrową estymację gęstości wyników (KDE – Kernel Density Estimation).

## Mapy ciepła to swego rodzaju macierze obrazujące wynik implementacji algorytmu korelacji Pearsona i liczonych przez niego współczynników. Współczynnik korelacji Pearsona przyjmuje wartości z zakresu <-1;1>. Wartość = „1” sygnalizuje silną korelację pozytywną pomiędzy analizowanymi wartościami natomiast wartość = „-1” silną korelację negatywną. Wartość środkowa „0” sygnalizuje natomiast brak jakiejkolwiek korelacji pomiędzy badanymi wartościami zmiennych. Podczas tworzenia matrycy korelacji identyfikowane są zależności pomiędzy zmiennymi numerycznymi opisującymi transakcje. W tym konkretnym przykładzie poszukiwana relacja pomiędzy zmienną decyzyjną „Class”, a pozostałymi parametrami opisującymi transakcje w taki sposób, by zidentyfikować te zmienne, które mają największy wpływ na estymację wartości parametru „Class”.

## C:\Users\Piotr Laskowski\AppData\Local\Microsoft\Windows\INetCache\Content.Word\EDA_6.png

Rysunek 7: Macierz korelacji Pearsona zwizualizowana w postaci mapy ciepła