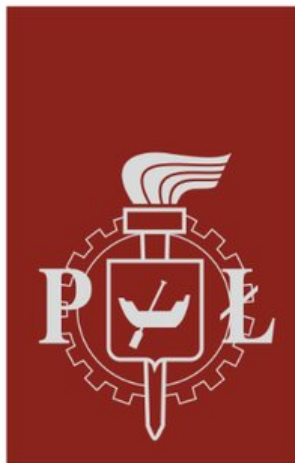


Politechnika Łódzka

Wydział Elektrotechniki Elektroniki Informatyki i Automatyki



sem, zimowy, r ak. 2024/2025

Sprawozdanie z projektu BigData „Predykcja cen samochodów używanych”



Mateusz Grzybek 240678

Kamil Młynarczyk 240757

15 grudnia 2024

Spis treści

1	Wstęp	2
1.1	Założenia projektowe	2
1.2	Wykorzystane technologie	2
2	Aplikacja kliencka	3
2.1	Opis	3
2.2	Widoki aplikacji	4
2.2.1	Strona	4
2.2.2	Okno z ceną	5
2.2.3	Okno z błędem	5
3	Komponent pośredniczący	6
3.1	Opis	6

Rozdział 1

Wstęp

1.1 Założenia projektowe

Celem projektu jest zaimplementowanie aplikacji webowej pozwalającej użytkownikom na predykcję ceny używanego samochodu na podstawie dostarczonego przez niego zestawu cech. Tematyka projektu daje możliwość wykorzystania różnorodnych technologii z dziedziny uczenia maszynowego, rozwoju aplikacji webowych, komunikacji pomiędzy serwisami, architektury oprogramowania oraz bierania i przetwarzania danych. W celu zrealizowania przewidywanych funkcjonalności, aplikacja została podzielona na cztery komponenty, każdy z nich odpowiedzialny za realizację innego aspektu aplikacji.

1.2 Wykorzystane technologie

- Java — Obiektowy język programowania.
- SpringBoot — Framework dla języka Java nastawiony na wytwarzanie aplikacji webowych i mikroservisów
- Gradle — Narzędzie do automatyzacji budowania projektów.
- React — Framework JavaScript do tworzenia interfejsów użytkownika w oparciu o komponenty.
- Docker — Narzędzie do tworzenia, uruchamiania i zarządzania aplikacjami w izolowanych środowiskach zwanych kontenerami.
- Docker Compose — Narzędzie usprawniające zarządzanie wieloma kontenerami jednocześnie.
- Python — Język skryptowy.
- Apache Spark — Framework do sprawnego przetwarzania zbiorów danych w pamięci.
- Apache SparkML — Moduł Apache Spark przeznaczony do uczenia maszynowego.
- Apache Kafka — Platforma przetwarzania danych w czasie rzeczywistym.
- Apache Zookeeper — Usługa koordynacyjna systemów rozproszonych.

Rozdział 2

Aplikacja kliencka

2.1 Opis

Aplikacja kliencka stanowi pojedynczą stronę dostępną za pośrednictwem przeglądarki, udostępnianą pod adresem **localhost**¹, na porcie **9091**. Strona zawiera informacje związane z aplikacją oraz pola do wprowadzania wartości, na podstawie których następnie dokonywana jest predykcja ceny samochodu. Aplikacja łączy się z komponentem middleware za pośrednictwem protokołu **HTTP**² w architekturze **REST**³.

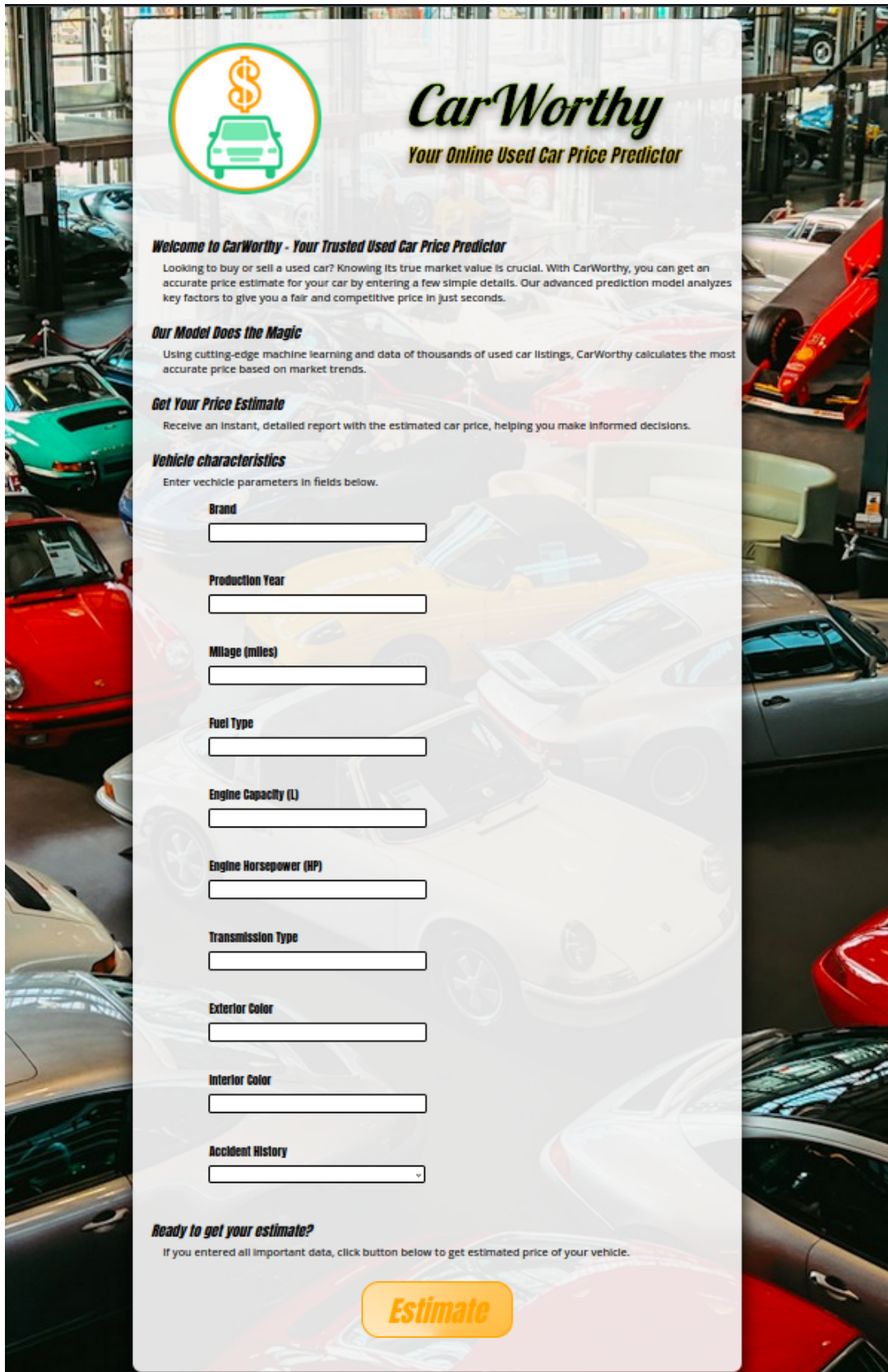
¹loopback address — adres pętli zwrotnej, który jest wykorzystywany do komunikacji urządzenia z samym sobą.

²HyperText Transfer Protocol — protokół komunikacyjny używany do przesyłania danych w sieci.

³Representational State Transfer — architektura komunikacji oparta o protokół HTTP definiujący sposoby identyfikacji i manipulacji zasobami za pomocą zapytań HTTP.

2.2 Widoki aplikacji

2.2.1 Strona



The screenshot displays the CarWorthy website, which is an online used car price predictor. The interface is set against a background image of a car dealership. At the top left, there is a logo featuring a green car icon with a dollar sign above it, enclosed in a green circle. To the right of the logo, the text "CarWorthy" is written in a large, stylized font, with "Your Online Used Car Price Predictor" in a smaller font below it.

Below the header, there is a welcome message: "Welcome to CarWorthy - Your Trusted Used Car Price Predictor". This is followed by a paragraph explaining the service: "Looking to buy or sell a used car? Knowing its true market value is crucial. With CarWorthy, you can get an accurate price estimate for your car by entering a few simple details. Our advanced prediction model analyzes key factors to give you a fair and competitive price in just seconds."

Next, there is a section titled "Our Model Does the Magic" with a paragraph: "Using cutting-edge machine learning and data of thousands of used car listings, CarWorthy calculates the most accurate price based on market trends."

Below this, there is a section titled "Get Your Price Estimate" with a paragraph: "Receive an instant, detailed report with the estimated car price, helping you make informed decisions."

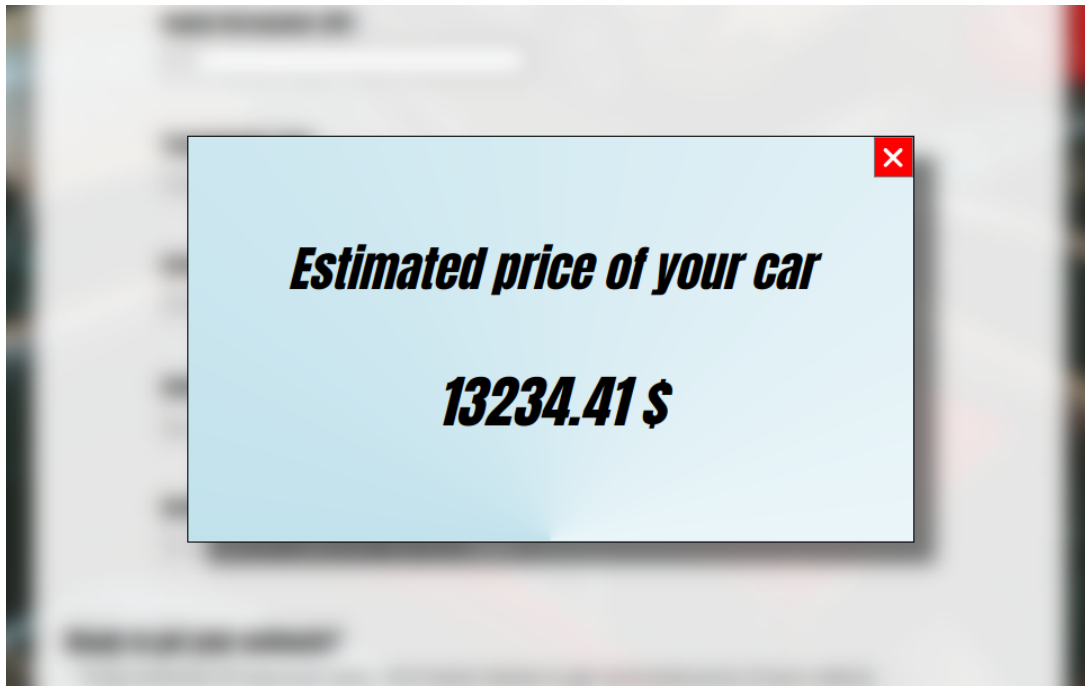
The main part of the form is titled "Vehicle characteristics" and includes the instruction "Enter vehicle parameters in fields below." There are ten input fields, each with a label and a text box:

- Brand
- Production Year
- Milage (miles)
- Fuel Type
- Engine Capacity (L)
- Engine Horsepower (HP)
- Transmission Type
- Exterior Color
- Interior Color
- Accident History

At the bottom of the form, there is a section titled "Ready to get your estimate?" with a paragraph: "If you entered all important data, click button below to get estimated price of your vehicle." Below this text is a large, orange, rounded rectangular button with the word "Estimate" in white text.

Rysunek 2.1: Widok strony

2.2.2 Okno z ceną



Rysunek 2.2: Widok okna z ceną

2.2.3 Okno z błędem



Rysunek 2.3: Widok okna z błędem

Rozdział 3

Komponent pośredniczący

3.1 Opis

Komponent pośredniczący pełni rolę pośrednika pomiędzy aplikacją kliencką i serwisem predykcyjnym. Otrzymywane od **frontendu**¹ dane w formie **JSON**² są w tym komponencie przetwarzane na wiadomości w formacie odpowiadającym wejściu modelu, z uwzględnieniem procesu **kodowania liczbowego**³ pól. Otrzymane w tym procesie wiadomości zapisywane są na temat wejściowy Kafki. Pośrednik jest również odpowiedzialny za odczytywanie danych z tematu wyjściowego i przekazywanie uzyskanych z nich informacji do klienta.

¹Część aplikacji, z którą użytkownik wchodzi w bezpośrednią interakcję, w tym wszystko co widzi oraz elementy wizualne i interaktywne.

²JavaScript Object Notation — format danych zapewniający kompaktowe rozmiary i jest czytelny dla ludzi i maszyn.

³Technika zamiany wartości danych tekstowych na wartości liczbowe, poprzez przypisanie unikalnej liczby każdej unikalnej wartości tekstowej.