## POLITECHNIKA WROCŁAWSKA WYDZIAŁ ELEKTRONIKI

# Projektowanie systemów z dostępem w języku naturalnym

# Chatbot wspomagający obsługę klienta w sklepie internetowym

AUTORZY: PROWADZĄCY ZAJĘCIA:

Krystian Lema 218453 dr inż. Dariusz Banasiak, W4/K9

Bartosz Michalak 218481

OCENA PRACY:

## Spis treści

1.	Wst	tęp	3
			3
			3
			3
			4
2.	Imp	olementacja	5
	2.1.	Konfiguracja środowiska	5
	2.2.	Konfiguracja biblioteki ChatterBot	5
			5
			6
			7
			8
3.	Test	towanie	9
	3.1.	Podstawowe testy dialogowe	9
			9
		3.1.2. Pożegnania	0
	3.2.	Testy zestawów pytań	1
		3.2.1. Pytanie o formę dostawy	1
		3.2.2. Pytanie o status przesyłki	3
	3.3.	Zwiększanie efektywności bota poprzez zwiększenie bazy wiedzy 1	5
4.	Wn	ioski	6
		Ogólne wnioski na temat działania chatbota	
		Wpływ bazy wiedzy na skuteczność działania	
Lii	terati	ıra	7

## Wstęp

#### 1.1. Cel projektu

Celem projektu jest stworzenie chatbota, którego zadaniem jest wsparcie klienta w sklepie internetowym ze sprzętem komputerowym. Komunikacja ma odbywać się w trybie tekstowym. Zadaniem chatbota jest rozpoznanie pytania zadanego przez klienta sklepu internetowego oraz znalezienie odpowiedzi w bazie wiedzy. Wykorzystano również technikę uczenia maszynowego do poprawienia skuteczności chatbota.

#### 1.2. Założenia projektowe

Chatbot ma za zadanie pomagać klientom w następujących tematach:

- Obsługa konta użytkownika (zmiana hasła, przypomnienie hasła, bezpieczeństwo danych)
- Dostawa (czas, koszt, rodzaje)
- Zasady zwrotów (ile czasu na zwrot, koszt zwrotu, sposób)
- Dostępne metody płatności
- Śledzenie statusu zamówienia
- Zasady reklamacji
- Aktualne promocje

#### 1.3. Użyte technologie

Chatbot został napisany w języku programowania Python z użyciem biblioteki dedykowanej do tworzenia chatbotów ChatterBot. ChatterBot to silnik do tworzenia dialogów, który używa bazy wiedzy w postaci plików tekstowych oraz uczenia maszynowego do zwiększenia skuteczności znajdywania odpowiednich odpowiedzi.

#### 1.4. Zasada działania chatbota

#### Schemat działania:

- 1. Odczytanie danych
- 2. Przetwarzanie surowego tekstu
- 3. Tokenizacja tekstu
- 4. Usunięcie zbędnych znaków
- 5. Usunięcie mało znaczących słów
- 6. Sprowadzenie słów do podstawowych form (lematyzacja)
- 7. Parsowanie zależności
- 8. Generowanie odpowiedzi na podstawie bazy wiedzy

## Implementacja

#### 2.1. Konfiguracja środowiska

Do stworzenia bota użyto języka Python w wersji 3. W tym celu zainstalowano najnowszą wersję, którą można znaleźć na oficjalnej stronie: https://www.python.org/. Aby ułatwić pracę w środowisku, dodatkowo zainstalowano managera paczek do Python'a - PIP. Instrukcje dotyczące manualnej instalacji można znaleźć na oficjalnej stronie: https://pypi.org/project/pip/.

#### 2.2. Konfiguracja biblioteki ChatterBot

Instalacja biblioteki ChatterBot jest bardzo prosta przy użyciu PIP. Całą bibliotekę można zainstalować jedną komendą:

pip install chatterbot

Zainstalowaną bibliotekę można sprawdzić przy pomocy komendy:

python -m chatterbot -version

Terminal powinien zwrócić aktualną wersję.

#### 2.3. Implementacja chatbota

Pierwszym elementem programu było stworzenie obiektu chatbota. Wykonano to przy pomocy obiektu ChatBot z biblioteki ChatterBot. Podczas tworzenia obiektu chatbota można zainicjalizować go kilkoma parametrami.

Rys. 2.1: Tworzenie obiektu bota - kod źródłowy.

Pierwszym parametrem jest nazwa chatbota. Kolejnym jest lista adapterów logicznych, których można podać inicjując bota. Nie są one obowiązkowe, ale bez ich określenia, stworzony bot będzie posiadał adaptery domyślne. Aby mieć większą kontrolę nad chatbotem, podano kilka adapterów:

- MathematicalEvaluation dzięki temu adapterowi bot jest w stanie odpowiadać na pytania rozwiązując proste operacje matematyczne. Np. na pytanie 2 + 2, powinien zwrócić odpowiedź 4, która jest prawidłowym wynikiem dodawania.
- BestMatch adapter logiczny, który przy wybieraniu odpowiedzi wybiera tą która najlepiej pasuje do pytania lub po prostu jest tej odpowiedzi najbardziej pewny. Porównywanie odpowiedzi odbywa się przy użyciu algorytmu dogległości Levenstein'a (linia 20 kodu), a z pośród znalezionych odpowiedzi bot wybiera tą, która wystąpiła najczęsciej (linia 21 kodu).
- LowConfidenceAdapter adapter odpowiadający za działanie w przypadku gdy bot nie jest pewny odpowiedzi na zadane pytanie. Adapter ten pozwala ustawić próg poniżej, którego bot nie jest w stanie odpowiedzieć i zwróci tekst informujący o tym np. *Przepraszam ale nie jestem w stanie odpowiedzieć na to pytanie*.

Ostatnim parametrem są filtry. Tutaj dodano filtr RepetitiveResponseFilter, który zapobiega powtarzaniu tych samych odpowiedzi przez bota.

#### **2.3.1.** Uczenie

Kolejnym etapem jest uczenie bota. Uczenia można dokonać przy pomocy plików tekstowych, zawierających pytania i odpowiedzi w odbpowiednim schemacie. W tym celu stworzono folder, w którym umieszczono wszystkie pliki uczące w formacie *txt*.

```
# Train bot from files
bot.set_trainer(ListTrainer)
for _file in os.listdir('train_files'):
    chats = open('train_files/' + _file, 'r').readlines()
    bot.train(chats)
```

Rys. 2.2: Uczenie bota - kod źródłowy.

Program wczytuje wszystkie pliki w zadanym folderze, otwiera je w trybie odczytu i czyta wszystkie linie. Bot uczy się z każdego pliku, zapamiętuje pytania i odpowiedzi dodając je do bazy danych.

#### 2.3.2. Baza wiedzy

W celu nauczenia bota odpowiedzi na pytania z różnych tematów przygotowano odpowiednie pliki w formacie tekstowym *txt*. Każdy plik odpowiada za inny temat dialogu. Przygotowano odpowiednio dane o tematyce:

dostawa - konto klienta
 płatności - powitania
 pożegnania - promocje
 reklamy - zwroty

Format pliku to pytanie i odpowiedź w kolejnych liniach pliku.

```
Jakie są formy dostawy?
     Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej
     informacji znajdziesz na stronie http://sklep-internetowy.pl/dostawy
   Ile kosztuje dostawa?
   Koszt dostawy do paczkomatu to 8 zł, Pocztą Polską 10 zł, natomiast kurierem
    12 zł.
   Ile kosztuje kurier?
    Koszt dostawy kurierem to 12 zł.
    Ile kosztuje dostawa Poczta Polską?
    Koszt dostawy Pocztą Polską to 10 zł.
    Ile kosztuje dostawa do paczkomatu?
10
   Koszt dostawy do paczkomatu to 8 zł.
   Jaki jest czas dostawy?
   Czas dostawy zależy od formy doręczenia. Wszystkie paczki przygotujemy oraz
    nadajemy w dzień zamówienia.
    Dlaczego jeszcze nie ma mojej paczki?
    Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
    mailu z potwierdzeniem nadania wysyłki.
    Chciałbym sprawdzić status wysyłki.
    Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
    mailu z potwierdzeniem nadania wysyłki.
    Kiedy będzie moja paczka?
    Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
    mailu z potwierdzeniem nadania wysyłki.
    Jak mogę śledzić przesyłkę?
    Przesyłkę możesz śledzić za pomocą kodu śledzącego, który otrzymałeś w mailu
    z potwierdzeniem nadania wysyłki. Znajdziesz tam również link do strony, na
    której należy wpisać otrzymany kod śledzący.
    Kiedy otrzymam przesyłkę?
    W przypadku Poczty Polskiej czas doręczenia to 72h od momentu nadania, Kurier
    - 48h, Paczkomat - 24h. Aktualny status przesyłki możesz sprawdzić za pomocą
     kodu śledzącego, który otrzymałeś w mailu z potwierdzeniem nadania przesyłki.
    Jak odebrać przesyłkę z Paczkomatu?
    Dokładną instrukcję znajdziesz w mailu z potwerdzeniem nadania wysyłki.
    Czy przesyłka jest ubezpieczona?
    Tak, wszystkie przesyłki są ubezpieczona.
    Moja przesyłka jest uszkodzona.
    Proszę o zgłoszenie reklamacji za pomocą formularza, który znajdzie Pan na
    stronie http://sklep-interntowy/reklamacja-przesylki.
    Moje zamówienie jest niekompletne.
    Proszę o zgłoszenie reklamacji za pomocą formularza, który znajdzie Pan na
    stronie http://sklep-interntowy/reklamacja-przesylki.
    Paczka przyszła uszkodzona.
32 Proszę o zgłoszenie reklamacji za pomocą formularza, który znajdzie Pan na
    stronie http://sklep-interntowy/reklamacja-przesylki.
```

Rys. 2.3: Przykładowy plik z danymi - dostawa.txt.

#### 2.3.3. Dialog bota z użytkownikiem

```
continue_dialogue = True
# Start of the conversation
print('{0}: {1}'.format(bot_display_name, bot_greetings))
while continue_dialogue:
    # Get input from user
    request = input('{0}: '.format(user_display_name))
   # Check if conversation ends or continues
    continue_dialogue = True
   for bye word in bye words:
       if bye_word in request.lower():
            continue_dialogue = False
   # Display chatbot response
    if continue_dialogue:
        response = bot.get_response(request)
        print('{0}: {1} ({2})'.format(bot_display_name, response, response.confidence))
      print('{0}: {1}'.format(bot_display_name, bot_bye_answer))
```

Rys. 2.4: Dialog bota z użytkownikiem - kod źródłowy.

Na samym początku bot zaczyna dialog witając użytkownika i czekając na wprowadzenie tekstu. Dialog zapętlony jest aż do momentu kiedy użytkownik zakończy konwersacje jednym z definiowanych słów pożegnalnych (do widzenia, żegnam, koniec). Po wprowadzeniu tekstu przez użytkownika sprawdzane jest czy nie zakończył on konwersacji. Jeżeli tak, bot żegna się i program kończy działania. W przeciwnym wypadku bot analizuje wprowadzony przez użytkownika tekst i wybiera najlepszą odpowiedź. Gdy pewność odpowiedzi bota jest większa niż przyjęty próg 0.4, odpowiedź zostaje wyświetlona. Dla ułatwienia testowania, z każdą odpowiedzią bota na końcu wyświetlana jest jego pewność jako wartość z przedziału  $\{0,1\}$ .

```
Serwis sklepu internetowego: Witam! Jak mogę pomóc?

Użytkownik: Jaki jest czas dostawy?

Serwis sklepu internetowego: Czas dostawy zależy od formy doręczenia. Wszystkie paczki przygotujemy oraz nadajemy w dzień zamówienia. (1.0)

Użytkownik: Dostawa?

Serwis sklepu internetowego: Koszt dostawy do paczkomatu to 8 zł, Pocztą Polską 10 zł, natomiast kurierem 12 zł. (0.55)

Użytkownik: Dziękuję. Do widzenia.

Serwis sklepu internetowego: Do widzenia. Dziękujemy za skorzystanie z naszych usług i zapraszamy ponownie!
```

Rys. 2.5: Przykładowy dialog.

### **Testowanie**

#### 3.1. Podstawowe testy dialogowe

#### 3.1.1. Powitania

Jak wspomniano w poprzednim rozdziale, przygotowano zestaw pytań pozwalający na przywitanie się i pożegnanie się przez bota. Przeprowadzono testy sprawdzające, jak reaguje na powitania użytkownika:

```
Serwis sklepu internetowego: Witam! Jak mogę pomóc?
Użytkownik: Dzień dobry.
Serwis sklepu internetowego: Witam. W czym mogę pomóc? (1.0)
Użytkownik: Witaj.
Serwis sklepu internetowego: Witam. W czym mogę pomóc? (1.0)
Użytkownik: Witam.
Serwis sklepu internetowego: Witam. W czym mogę pomóc? (1.0)
Użytkownik: Cześć.
Serwis sklepu internetowego: Witam. W czym mogę pomóc? (1.0)
```

Rys. 3.1: Reakcja bota na powitania użytkownika.

Jak można zaobserwować, pewność odpowiedzi wynosi 100%. Wynika to z tego, że w bazie wiedzy znajdują się dokładnie te zwroty, których użyto podczas testów.

Kolejnym krokiem było sprawdzenie reakcji na powitania, które zawierają błędy literowe lub zamienioną kolejność wyrazów.

#### Użytkownik: Dzien dobry Serwis sklepu internetowego: Witaj. (0.87)

Rys. 3.2: Reakcja bota na powitanie użytkownika zawierające błąd literowy.

Można zaobserwować zmniejszenie pewności odpowiedzi o 13%.

Kolejnym eksperymentem było sprawdzenie, jak użycie znaków interpunkcyjnych wpływa na pewność odpowiedzi. Wszystkie sekwencje w bazie wiedzy zapisano ze znakami interpunkcyjnymi na końcu zdań oraz pytań.

```
Użytkownik: Dzien dobry
Serwis sklepu internetowego: Witaj. (0.87)
```

Rys. 3.3: Powitanie bez znaku interpunkcyjnego na końcu zdania.

```
Użytkownik: Dzien, dobry
Serwis sklepu internetowego: Witaj. (0.83)
```

Rys. 3.4: Powitanie ze znakiem interpunkcyjnym pomiędzy dwoma słowami.

Na rysunku 3.3. można zaobserwować obniżenie pewności o 13%. Na rysunku 3.4. widać obniżenie pewności o kolejne 4% w związku z dodaniem niewystępującego przecinka pomiędzy dwoma wyrazami.

Wartości te jednak należy uznać za akceptowalne, mając na uwadze, że minimalna pewność odpowiedzi została ustalona na 40%.

#### 3.1.2. Pożegnania

W bazie wiedzy zawierającej możliwe pożegnania użytkownika zdeklarowano krótkie, w większości składające się z jednego słowa zdania:

```
Dziękuję.
Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc?
Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc?
Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc?
Dziękuję.
Dziękuję za rozmowę, mam nadzieję że udało mi się rozwiać wszystkie wątpliwości. Czy mogę w czymś jeszcze pomóc?
Do widzenia.
Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc?
Dobranoc.
Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc?
Na razie.
Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc?
```

Rys. 3.5: Baza wiedzy dotycząca pożegnań.

Testując interakcję bota w przypadku pożegnania sprawdzono, jak na pewność odpowiedzi wpłynie dodanie dodatkowego słowa do wypopowiedzi użytkownika.

```
Użytkownik: Dziękuję.
Serwis sklepu internetowego: Dziękuję za rozmowę, czy mogę w czymś jeszcze pomóc? (1.0)
Użytkownik: Dziękuję, pozdrawiam.
Serwis sklepu internetowego: Dziękuję za rozmowę, mam nadzieję że udało mi się rozwiać wszystkie wątpliwości. Czy mogę w
czynś jeszcze pomóc? (0.6)
```

Rys. 3.6: Pojawienie się dodatkowego słowa w wypowiedzi użytkownika.

Na rysunku 3.6. można zaobserwować obniżenie pewności o 40% w związku z dodaniem jednego słowa do pożegnania.

Kolejnym krokiem było dodanie kolejnego słowa do wypowiedzi użytkownika i sprawdzenie, jak bot poradzi sobie z frazą różniącą się od bazowej o 66% (3 wyrazy zamiast jednego).

```
Jżytkownik: Dziękuję, pozdrawiam serdecznie.
Serwis sklepu internetowego: Jak się masz? (0.48)
```

Rys. 3.7: Pojawienie się dwóch dodatkowych słów w wypowiedzi użytkownika.

Zwrócona przez bota nie jest zgodna z oczekiwaną. Współczynnik pewności wynosi 48% - rozwiązaniem tego problemu mogłoby być m.in.:

- zwiększenie progu pewności, powyżej którego bot odpowiada na podstawie bazy wiedzy;
- rozszerzenie bazy wiedzy.

#### 3.2. Testy zestawów pytań

Powitania oraz pożegnania to ważna część funkcjonalności bota, jednak nie najważniejsza. Główną funkcjonalność stanowi możliwość pomocy klientowi w podstawych sprawach związanych z zakupami w sklepie internetowym. Kolejnym krokiem było więc przetestowanie wybranych zestawów pytań i odpowiedzi.

#### 3.2.1. Pytanie o formę dostawy

Pierwszym pytaniem, jakie zostało poddane testom, było pytanie o dostępne formy dostawy. W bazie wiedzy zostało ono zapisane w następujący sposób:

```
Jakie są formy dostawy?
Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdziesz na stronie http://sklep-internetowy.pl/dostawy
```

Rys. 3.8: Pytanie o formy dostawy zapisane w bazie wiedzy.

Sprawdzono następujące przypadki:

• pytanie zgodne w 100% ze zdefiniowanym w bazie wiedzy.

```
Użytkownik: Jakie są formy dostawy?
Serwis sklepu internetowego: Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdzie
sz na stronie http://sklep-internetowy.pl/dostawy (1.0)
```

Rys. 3.9: Odpowiedź bota na pytanie zgodne w 100% ze zdefiniowanym w bazie wiedzy.

• pytanie bez znaku interpunkcyjnego

```
Użytkownik: Jakie są formy dostawy
Serwis sklepu internetowego: Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdzie
sz na stronie http://sklep-internetowy.pl/dostawy (0.98)
```

Rys. 3.10: Odpowiedź bota na pytanie bez znaku interpunkcyjnego.

Warto zwrócić uwagę, że pewność obniżyła się jedynie o 2%, podczas gdy w przypadku braku znaku interpunkcyjnego dla zdania składającego się z dwóch słów było to 13%.

• pytanie zawierające synonim

```
Użytkownik: Jakie są możliwości dostawy?
Serwis sklepu internetowego: Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdzie
sz na stronie http://sklep-internetowy.pl/dostawy (0.75)
```

Rys. 3.11: Odpowiedź bota na pytanie zawierające synonim 'możliwości'.

```
Użytkownik: Jakie są opcje dostawy?
Serwis sklepu internetowego: Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdzie
sz na stronie http://sklep-internetowy.pl/dostawy (0.83)
```

Rys. 3.12: Odpowiedź bota na pytanie zawierające synonim 'opcje'.

Warto zwrócić uwagę, że pewność boa była inna dla różnych synonimów. Wynika to z liczby znaków poszczególnych synonimów.

• pytanie zawierające dodatkowe słowo

```
Użytkownik: Jakie są dostępne formy dostawy?
Serwis sklepu internetowego: Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdzie
sz na stronie http://sklep-internetowy.pl/dostawy (0.84)
```

Rys. 3.13: Odpowiedź bota na pytanie zawierające dodatkowe słowo.

• pytanie zawierające dodatkowe słowo oraz synonim

```
Użytkownik: Jakie są dostępne formy dostawy?
Serwis sklepu internetowego: Paczki dostarczamy Paczkomatem, Pocztą Polską oraz Kurierem DPD. Więcej informacji znajdzie
sz na stronie http://sklep-internetowy.pl/dostawy (0.84)
```

Rys. 3.14: Odpowiedź bota na pytanie zawierające dodatkowe słowo.

#### 3.2.2. Pytanie o status przesyłki.

Badany przypadek użycia został w bazie wiedzy zdefiniowany w następujący sposób:

```
Dlaczego jeszcze nie ma mojej paczki?
Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w mailu z potwierdzeniem nadania wysyłki.
```

Rys. 3.15: Definicja pytania w bazie wiedzy

Testom poddano te same sytuacje, co w przypadku poprzedniego pytania. Tym razem jednak badane pytanie zawierało więcej słów.

• pytanie zgodne w 100% ze zdefiniowanym w bazie wiedzy.

```
Serwis sklepu internetowego: Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
mailu z potw<u>i</u>erdzeniem nadania wysyłki. (1.0)
```

Rys. 3.16: Odpowiedź bota na pytanie zgodne w 100% ze zdefiniowanym w bazie wiedzy.

• pytanie bez znaku interpunkcyjnego

```
Jżytkownik: Dlaczego jeszcze nie ma mojej paczki
Serwis sklepu internetowego: Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
mailu z potw<u>i</u>erdzeniem nadania wysyłki. (0.99)
```

Rys. 3.17: Odpowiedź bota na pytanie bez znaku interpunkcyjnego.

Po raz kolejny można zaobserwować, że większa liczba słów ma pozytywny wpływ ma pewność odpowiedzi bota w sytuacji pominięcia znaku interpunkcyjnego.

• pytanie zawierające synonim

```
Użytkownik: Dlaczego jeszcze nie ma mojej przesyłki?
Serwis sklepu internetowego: Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
mailu z potwierdzeniem nadania wysyłki. (0.91)
```

Rys. 3.18: Odpowiedź bota na pytanie zawierające synonim 'przesyłki'.

Pewność bota obniżyła się jedynie o 9%.

• pytanie zawierające dodatkowe słowo

```
Jżytkownik: Dlaczego jeszcze nie ma mojej zamówionej paczki?
Serwis sklepu internetowego: Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
mailu z potw<u>i</u>erdzeniem nadania wysyłki. (0.87)
```

Rys. 3.19: Odpowiedź bota na pytanie zawierające dodatkowe słowo.

• pytanie zawierające dodatkowe słowo oraz synonim

```
Jżytkownik: Dlaczego jeszcze nie ma mojej zamówionej przesyłki?
Serwis sklepu internetowego: Status paczki możesz sprawdzić za pomocą kodu śledzącego znajdującego się w
mailu z potw<u>i</u>erdzeniem nadania wysyłki. (0.77)
```

Rys. 3.20: Odpowiedź bota na pytanie zawierające dodatkowe słowo oraz synonim

3.3. Zwiększanie efektywności bota poprzez zwiększenie bazy wiedz
Zwiększanie efektywności bota poprzez zwiększenie bazy wiedzy

## Wnioski

- 4.1. Ogólne wnioski na temat działania chatbota
- 4.2. Wpływ bazy wiedzy na skuteczność działania

## Literatura

[1] Dokumentacja ChatterBot'a

https://chatterbot.readthedocs.io/en/stable/chatterbot.html

dostęp: 08.01.2019

[2] Dokumentacja Python'a

https://docs.python.org/3/

dostęp: 08.01.2019