

# Projekt SAG/WEDT

Temat:

Przewidywanie ratingu/oceny na podstawie opisu tekstowego (np. recenzje filmów, produktów etc.)

Autorzy:

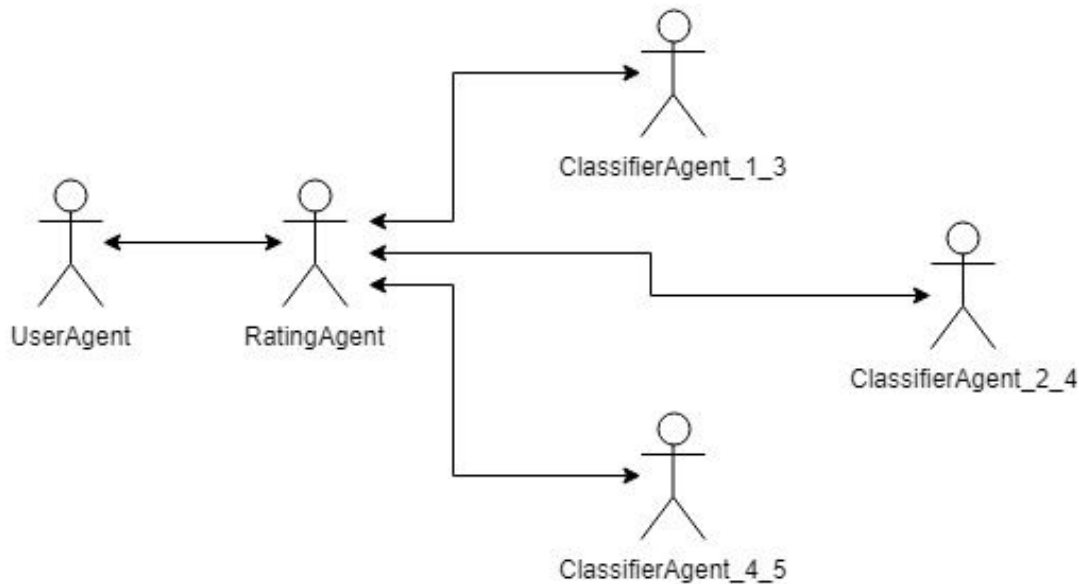
Dziewulska Katarzyna

Paczuski Paweł

Wereszczyński Łukasz

## 1. Architektura systemu

W ramach projektu postanowiliśmy zaimplementować następującą architekturę systemu agentowego:



Działanie systemu wyglądałoby następująco:

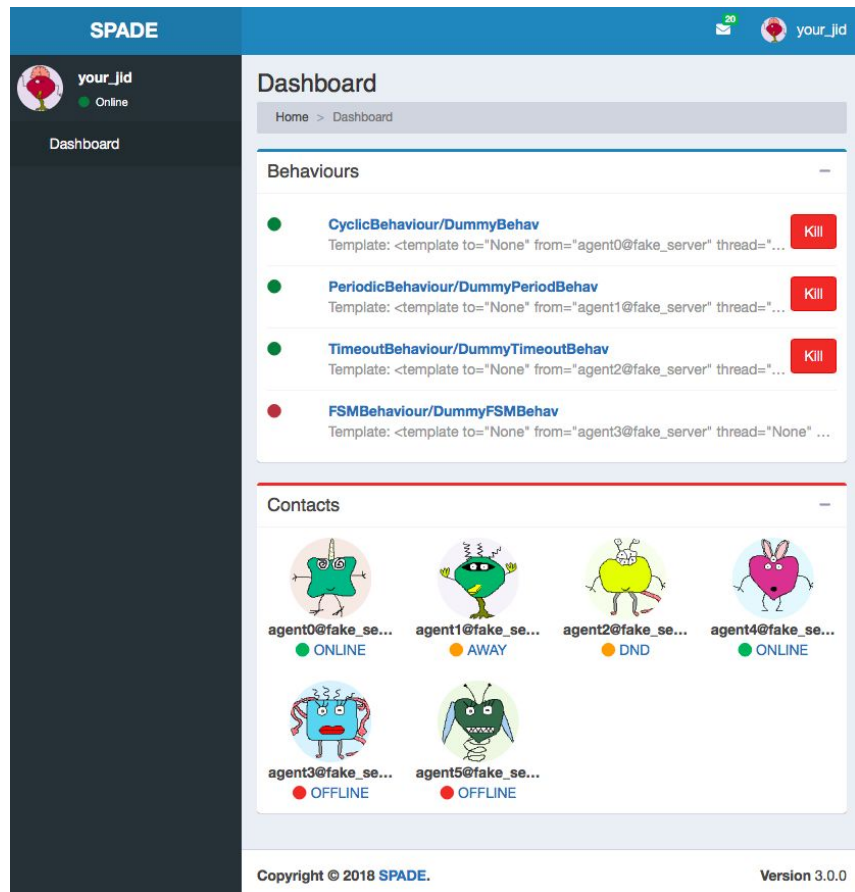
1. Użytkownik przekazuje do systemu recenzję lub opinię na temat filmu/ produktu którego ocenę chciałby poznać. Jest on reprezentowany przez **UserAgent'a**.
2. Taka wiadomość zostaje przekazana do **RatingAgent'a**, który wysyła wiadomość do **ClassifierAgent'ów** z prośbą o ocenę danej recenzji / opinii.
3. Następnie, **ClassifierAgent'ci** przystępują do oceniania przekazanej im recenzji. Jest czterech **ClassifierAgent'ów**, którzy są specjalistami w rozpoznaniu odpowiednich ocen (1-4).
4. **ClassifierAgent'ci** dokonują oceny na podstawie wytrenowanego klasyfikatora. Trenowanie będzie przeprowadzone na części datasetu składającego się z konkretnej grupy ocen, wejściem będzie w postaci liczności **Bag of words** oraz postaci **n-gram'owej** recenzji/opinii.
  - a. Wstępnie będzie to jeden z najlepiej wytrenowanych klasyfikatorów:
    - i. Naiwny klasyfikator Bayes'a
    - ii. Drzewo decyzyjne
    - iii. SVM

5. W kolejnej fazie, agenci zaczynają uzgadniać ocenę końcową dla przekazanej im recenzji na podstawie średniej ocen zwróconych przez **ClassifierAgentów**.
6. Ocena końcowa zwracana jest użytkownikowi.
7. W przypadku, gdy jeden z Classifier agentów nie odpowie średnia wyznaczana jest na podstawie pozostałych ocen. Jeśli żaden z agentów nie odpowie system sygnalizuje o błędnym działaniu programu. Jeśli **RatingAgent** nie odpowie, wtedy użytkownik nie dostanie żadnej odpowiedzi.

## 2. Użyta platforma agentowa

SPADE jest prostą platformą agentową napisaną w pythonie. Konceptyjnie agenty traktowane są jak użytkownicy grupowego czatu. Mogą się do niego logować, mają dostęp do listy innych użytkowników "online", mogą wymieniać między sobą wiadomości. Komunikacja między agentami oparta jest na protokole XMPP. Agenty mogą mieć definiowane zachowania różnych typów, np. powtarzające się cyklicznie, implementacja maszyny stanów skończonych.

Podobieństwa do platformy wymiany wiadomości w czasie rzeczywistym są dodatkowo podkreślane przez twórców za pomocą web-interface służącego do debugingu. Wygląda jak typowy chat room, w którym dostępne agenty wymieniają między sobą wiadomości. Można przeglądać historię wiadomości agentów a także wstrzykiwać własne wiadomości do kanałów wymiany informacji.



### 3. Dataset

Scale dataset dostępny pod adresem

[https://www.cs.cornell.edu/people/pabo/movie-review-data/scale\\_data.tar.gz](https://www.cs.cornell.edu/people/pabo/movie-review-data/scale_data.tar.gz)

zawiera treści recenzji filmów wraz z ocenami wystawionymi przez recenzentów w skalach 0-2 (trzystopniowej) oraz 0-3 (czterostopniowej).

Teksty recenzji zostały lekko obrobione: wszystkie słowa są w formie lower-case, zauważyć także można dodanie spacji przed znakami interpunkcji w celu łatwiejszego podziału tekstu na słowa (tokenizacja). Autorzy data-setu starali się usunąć z tekstów dosłowne wartości liczbowe ocen.

## 4 Wyniki skuteczności klasyfikatora

### 4.1 Model recenzji w postaci *bag of words*

- Naiwny klasyfikator Bayes'a

miara	klasyfikator 1	klasyfikator 2	klasyfikator 3	klasyfikator 4
<i>recall</i>	0.55	0.55	0.58	0.57
<i>accuracy</i>	0.86	0.61	0.64	0.84
<i>precision</i>	0.51	0.51	0.45	0.49
<i>fallout</i>	0.09	0.35	0.34	0.1

- Drzewo decyzyjne

miara	klasyfikator 1	klasyfikator 2	klasyfikator 3	klasyfikator 4
<i>recall</i>	0.23	0.48	0.44	0.26
<i>accuracy</i>	0.77	0.57	0.58	0.76
<i>precision</i>	0.22	0.47	0.38	0.23
<i>fallout</i>	0.13	0.37	0.35	0.15

### 4.2 Model recenzji w postaci bigramów

- Naiwny klasyfikator Bayes'a

miara	klasyfikator 1	klasyfikator 2	klasyfikator 3	klasyfikator 4
<i>recall</i>	0.22	0.44	0.31	0.14
<i>accuracy</i>	0.84	0.59	0.60	0.78
<i>precision</i>	0.41	0.51	0.38	0.29
<i>fallout</i>	0.06	0.30	0.25	0.07

- Drzewo decyzyjne

miara	klasyfikator 1	klasyfikator 2	klasyfikator 3	klasyfikator 4
<i>recall</i>	0.13	0.51	0.41	0.16
<i>accuracy</i>	0.79	0.56	0.56	0.78
<i>precision</i>	0.20	0.47	0.36	0.22
<i>fallout</i>	0.09	0.40	0.37	0.11