Kraków, 15.03.2023

Dr hab. inż. Piotr Nawrocki, prof. AGH
Instytut Informatyki
Wydział Informatyki, Elektroniki i Telekomunikacji
Akademia Górniczo-Hutnicza w Krakowie
Al. Mickiewicza 30
30-059 Kraków
e-mail: piotr.nawrocki@agh.edu.pl

RECENZJA
rozprawy doktorskiej mgr. inż. Łukasza Pawlika pt.
„Metody zwiększenia skuteczności wirtualnych konsultantów poprzez minimalizację niepoprawnie rozpoznawanych intencji z wypowiedzi klientów"

1. Oświadczenia

1. Nie mam wspólnych prac z doktorantem mgr. inż. Łukaszem Pawlikiem, nie prowadziłem z nim wspólnych prac badawczych, nie oceniałem jego dorobku wydawniczego ani naukowego i nie pozostajemy w żadnej zależności służbowej.
2. Niniejsza recenzja została opracowana na podstawie pisma prof. dr hab. inż. Andrzeja Polańskiego, Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej z dnia 10.02.2023 r.
3. Recenzja jest opracowana zgodnie z wymaganiami określonymi w ustawie „Prawo o szkolnictwie wyższym i nauce" z dnia 20 lipca 2018 r., Dz.U. 2022 poz. 574, z późn. zm.

2. Ocena czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego

Rozprawa mgr. inż. Łukasza Pawlika dotyczy poprawy skuteczności wirtualnych konsultantów poprzez minimalizację niepoprawnie rozpoznawanych intencji z wypowiedzi klientów. Zagadnienia poruszane w pracy dotyczą obszarów informatyki i lingwistyki. Obecnie coraz popularniejsze stają się systemy pozwalające na automatyczną obsługę klienta. Oprócz klasycznych rozwiązań wykorzystujących rozmowy telefoniczne coraz częściej wykorzystywane są nowoczesne narzędzia takie jak voiceboty i chatboty.

Wdrożenie voicebota wymaga opracowania efektywnych mechanizmów automatycznego rozpoznawania mowy ASR (*automatic speech recognition*). Skuteczność takich mechanizmów jest bardzo wysoka w warunkach studyjnych, gdzie jakość sygnału dźwiękowego jest wysoka. Jednakże w warunkach Contact Center osiągnięcie wysokiej skuteczności rozpoznawania mowy odpowiedniej dla wirtualnego konsultanta stanowi duże wyzwanie w szczególności dla mniej popularnych języków, do jakich należy język polski. Innym bardzo ważnym elementem umożliwiającym efektywne działanie tego typu narzędzi jest moduł rozumienia języka naturalnego NLU (*natural language understanding*), którego kluczową cechą jest rozpoznawanie intencji na podstawie fraz tekstowych z wypowiedzi klientów. Równocześnie dla modułów ASR i NLU istotną rolę odgrywają wykorzystane korpusy językowe a ich skuteczność może być różna dla poszczególnych modeli językowych.

Zagadnienia rozważane w rozprawie doktorskiej mgr. inż. Łukasza Pawlika mogą być przedmiotem badań na gruncie dyscypliny „Informatyka Techniczna i Telekomunikacja".

Mgr inż. Łukasz Pawlik sformułował następującą tezę swojej rozprawy doktorskiej:

Możliwa jest poprawa skuteczności wirtualnych konsultantów w Contact Center poprzez:

1. *minimalizację niepoprawnie rozpoznawanych wypowiedzi klientów z zastosowaniem preprocessingu i postprocessingu danych systemu ASR,*
2. *rozpoznawanie ukrytych intencji klientów na podstawie emocji zawartych w ich wypowiedziach.*

Oceniam pozytywnie tezę rozprawy. Jednak można rozważyć czy nie warto byłoby bardziej uściślić w tezie pojęcia preprocessingu i postprocessingu wskazując wykorzystane w tych procesach mechanizmy.

Mgr inż. Łukasz Pawlik rozwiązał postawiony problem opracowując oryginalny system zwiększający skuteczność wirtualnego konsultanta w Contact Center zawierający moduły preprocessingu i postprocessingu oraz metody poprawy jakości transkrypcji i rozpoznawania ukrytych intencji. Architektura opracowanego systemu zawiera następujące moduły:

- moduł przygotowania danych uczących i testujących,
- moduł transkrypcji rozmów głosowych,
- moduł rozpoznawania emocji,
- moduł uczenia modelu NLU,
- moduł rozpoznawania ukrytych intencji.

Na podstawie przedstawionej architektury (Rysunek 7.1) można prześledzić działanie systemu. Nie jest jednak w pełni jasne czy na przykład moduł przygotowania danych uczących lub moduł uczenia (w tym model NLU) są autorskim rozwiązaniami mgr. inż. Łukasza Pawlika czy też zgodnie z opisem są nimi tylko wyróżnione elementy. Wydaje się że wszystkie elementy architektury były opracowane przez mgr. inż. Łukasza Pawlika a te wyróżnione są oryginalne w stosunku do innych systemów tego typu. Brakuje jednak jasnego wytłumaczenia tej kwestii.

W celu opracowania poszczególnych modułów systemu mgr inż. Łukasz Pawlik

zastosował szereg podejść naukowych, które odpowiednio zaadaptował w tym analizę sygnału dźwiękowego oraz NLU. Oryginalność opracowanego przez mgr. inż. Łukasza Pawlika rozwiązania polega na integracji wielu podejść, które zaadaptował do rozwiązania złożonego problemu poprawy skuteczności wirtualnego konsultanta. Na gruncie nauk technicznych jest to właściwe podejście.

W celu wykazania poprawności i skuteczności opracowanego rozwiązania mgr inż. Łukasz Pawlik zaimplementował i przetestował system zwiększający skuteczność wirtualnego konsultanta w Contact Center. Trzeba docenić wykorzystanie w testach rzeczywistych danych. Także bardzo wartościowe są wyniki analizy systemów automatycznego rozpoznawania mowy (Tabela 2.1) wraz z oceną ich skuteczności z wykorzystaniem metryki WER (*word error rate*).

Rozprawa jest czytelnie zredagowana i napisana poprawnym językiem. Jednak jednym z jej mankamentów jest bibliografia w której pojawiają się błędne pozycje w tym odnośniki do cyklu prac z Google Scholar ([72][73]), odniesienie do hasła z Wikipedii ([88]) oraz odniesienie do artykułu, który nie został jeszcze nigdzie opublikowany ([97]).

Mam też pewne uwagi do rozdziału 2. W podrozdziale 2.2 znajdują się rysunki 2.1 i 2.2 na których przedstawiono standardowe architektury voicebota i chatbota. Miejsce zamieszczenia tych rysunków sugeruje że są one opracowane na podstawie literatury ale brak jest odnośników do pozycji bibliograficznych. Także w tym podrozdziale brakuje analizy literatury pod kątem różnych architektur wirtualnych konsultantów co byłoby wskazane w kontekście tematu rozprawy.

3. **Ocena czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną mgr. inż. Łukasza Pawlika w dyscyplinie Informatyka Techniczna i Telekomunikacja**

Rozprawa doktorska mgr. inż. Łukasza Pawlika prezentuje jego szeroką wiedzę teoretyczną w różnych obszarach informatyki technicznej. Mgr inż. Łukasz Pawlik wykazał się głęboką wiedzą w zakresie systemów ASR oraz NLU. Ponadto wykazał się wiedzą z zakresu analizy lingwistycznej (wraz z rozpoznawaniem intencji) oraz uczenia maszynowego. Mgr inż. Łukasz Pawlik wykazał się również umiejętnościami praktycznymi implementując system zwiększający skuteczność wirtualnego konsultanta w Contact Center i przeprowadzając ocenę jego skuteczności.

4. **Ocena czy rozprawa doktorska wykazuje umiejętność samodzielnego prowadzenia pracy naukowej przez mgr. inż. Łukasza Pawlika**

Rozprawa doktorska mgr. inż. Łukasza Pawlika wykazuje jego umiejętność samodzielnego prowadzenia pracy naukowej w szczególności sformułowania problemu badawczego w obszarze informatyki technicznej i telekomunikacji, wypracowania oryginalnego rozwiązania postawionego problemu przy użyciu właściwych metod naukowych, wykazania poprawności opracowanego rozwiązania poprzez realizacje testów systemu oraz właściwej redakcji tekstu rozprawy.

5. **Konkluzja**

Biorąc pod uwagę aktualność tematyki doktoratu oraz osiągnięte wyniki badawcze

stwierdzam, że rozprawa doktorska mgr. inż. Łukasza Pawlika pomimo pewnych uwag krytycznych spełnia wymagania obowiązującej ustawy „Prawo o szkolnictwie wyższym i nauce" i wnoszę o jej dopuszczenie do publicznej obrony.