Informacja o wyniku oceny śródokresowej wraz z uzasadnieniem

doktoranta mgra inż. Mariusza Staronia kształcącego się w dyscyplinie automatyka, elekronika i elektrotechnika

Po zapoznaniu się z raportem z postępów w pracy naukowej oraz wysłuchaniu prezentacji oraz odpowiedzi na postawione pytania komisja ds. oceny śródokresowej wydała ocenę ***pozytywną***.

Szczegółowe uzasadnienie oceny:

Doktorant postawił sobie za główny cel klasyfikację sytuacji awaryjnych w sieci wodociągowej z wykorzystaniem metod sztucznej inteligencji. Do realizacji zadania planowane jest wykorzystanie głębokich sieci neuronowych typu splotowego. Do generowania danych zaproponowano wykorzystanie oprogramowania EPANET, natomiast implementacja głębokich sieci neuronowych realizowana będzie w środowisku inżynierskim Matlab z wykorzystaniem przybornika Deep Learning.

W pierwszym roku realizacji pracy doktorskiej zdefiniowano trzy zadania badawcze:
1. przegląd literatury, okres realizacji XI 2020 r. -- II 2021 r., kamień milowy: opracowany spis literatury,
2. budowa stanowiska laboratoryjnego, okres realizacji II 2021 r. -- III 2021r., kamień milowy: raport ze specyfikacją komputera,
3. przygotowanie pomiarów, IV 2021 r. -- VI 2021 r., kamień milowy: raport z wykonanymi przykładami.

W drugim roku realizacji pracy doktorskiej zdefiniowano dwa zadania badawcze:
4. przygotowanie pomiarów (ciąg dalszy), X 2021 r. -- I 2022 r., kamień milowy: raport z wizualizacji danych,
5. użycie głębokiej sieci do detekcji wycieków, I 2022 r. -- VI 2022 r., kamień milowy: raporty opisujące dobór danych uczących i walidacyjnych oraz dobór struktury sieci neuronowej.

Ocena terminowości i jakości wykonanych zadań
Ad. Zadanie 1. Prezentacja multimedialna zawiera tylko jeden slajd na temat przeglądu literatury obejmujący zestawienie analizowanych wątków. Niestety, nie zaprezentowano wykazu literatury na temat rozważanego tematu, czyli detekcji wycieków w sieciach wodociągowych. Brak jest informacji o tym jakie metody do tej pory były stosowane w tym zakresie oraz jakie są ich wady zalety. Na podstawie stanu wiedzy w rozważanym temacie powinna pojawić się motywacja zastosowania głębokich sieci neuronowych do rozwiązania postawionego celu pracy. Nie wiadomo również czym kierował się doktorant wybierając do realizacji badań środowisko inżynierskie Matlab. Obecnie do wykonywania tego typu eksperymentów wielu badaczy wykorzystuje język Python z bibliotekami TensorFlow i Keras,

które posiadają olbrzymie możliwości w zakresie metod uczenia maszynowego, a w szczególności uczenia głębokiego i są postrzegane jako efektywne rozwiązanie do złożonego obliczeniowo przetwarzania w Matlabie. Bez raportu zawierającego przegląd literatury trudno również określić terminowość wykonania zadania.

Ad. Zadanie 2. W prezentacji zawarto fotografie zbudowanego systemu komputerowego i podano specyfikację użytego sprzętu, czyli dane dotyczące jednostki przetwarzającej, pamięci RAM oraz procesora graficznego. Brakuje natomiast informacji o wykonanych testach wydajnościowych komputera (ang. *testbench*). Nie wiadomo jakie posiada on możliwości obliczeniowe w porównaniu do obecnie dostępnych komputerów powszechnego użytku. Jest to ważne, ponieważ doktorant wykorzystał procesor Intel Xeon X5690 z 2011 r. wykonany w litografii 32nm. Przy obecnie obserwowanym rozwoju techniki mikroprocesorowej można ten procesor uznać za przestarzały. W momencie zestawiania sprzętu komputerowego były już dostępne procesory Intel Xeon wykonane w technologii 14nm (np. Xeon 8356H z 8 rdzeniami). Podobnie wygląda sprawa karty graficznej. Premiera modelu ASUS Strix GTX 980 Ti miała miejsce w 2014 roku! Podsumowując, Intel Xeon to procesor zalecany do budowy serwerów, gdzie wymagana jest praca ciągła. Stąd nasuwa się pytanie, czy nie wystarczyłby procesor Inter Core i7 do realizacji badań. Na podstawie dostarczonych dokumentów nie można określić terminowości wykonania zadania, ale z dużą dozą pewności można uznać zadanie za wykonane.

Ad. Zadanie 3. Zadanie obejmowało zapoznanie się doktoranta z przybornikiem Deep Learning pakietu inżynierskiego Matlab. W prezentacji multimedialnej nie ma na ten temat żadnego odniesienia. Na slajdzie 13 jest jedynie informacja o wykonaniu zadania w 100\% jednak nie poparta żadnym wynikiem. Doktorant zawarł jedynie zrzut ekranu ze środowiska EPANET pokazujący na przykładową sieć wodociągową, aczkolwiek zadanie badawcze tego nie obejmowało. Na podstawie dostarczonej dokumentacji należy uznać zadanie jako niezrealizowane.

Ad. Zadanie 4. Zadanie obejmowało poznanie środowiska EPANET, wizualizację geodanych o sieci wodociągowej oraz wizualizację uzyskanych danych pomiarowej. Należy podkreślić, że doktorant poświęcił najwięcej miejsca w prezentacji multimedialnej omawiając te zagadnienia (12 z 30 slajdów). Analizując przedstawione informacje należy uznać zadanie za wykonane.

Ad. Zadanie 5. Zadanie obejmuje dwa działania: (i) przygotowanie danych do uczenia klasyfikatora, (ii) opracowanie struktury sieci neuronowej. W celu wykonania obu zadań wymagane jest przygotowanie formatu reprezentacji danych uczących. W prezentacji takich informacji nie ma. Co więcej, nie wiadomo nic na temat uszkodzeń jakie będą analizowane w ramach rozprawy doktorskiej. Czy system ma tylko wskazywać, że jakieś uszkodzenie w ogóle wystąpiło w sieci wodociągowej, czy ma informować także o typie/klasie uszkodzenia. Jest to ważne ze względu na kształt reprezentacji wyjściowej sieci neuronowej. W pierwszym przypadku to typowe zadanie dychotomii, w drugim już nie. Prezentacja nie podaje żadnych informacji na temat generowania danych uczących, ile wzorców zostało przygotowanych, jak wyglądał podział wzorców na uczące i testujące, czy w ramach zbioru uczącego był wydzielany zbiór walidacyjny oraz jak wprowadzano uszkodzenia i jak były reprezentowane w danych. Doktorant nie wyjaśnia również dlaczego zdecydował się na wykorzystanie sieci splotowych. Sieci splotowe wykazały swoją przydatność przy analizie obrazów graficznych. W przypadku

gdy dane są prezentowane w innej formie przydatne może być użycie klasycznego podejście wielowarstwowego albo sieci rekurencyjnych analizujących dane podane w postaci sekwencji czasowych, np. sieci GRU (ang. *Gated Recurrent Unit*) lub LSTM (ang. *Long Short-Term Memory*). Doktorant chce również zastosować algorytm Nguyena-Widrowa do inicjalizacji wag sieci neuronowej. Należy jednak podkreślić, że algorytm ten można wykorzystywać w przypadku stosowania ograniczonych funkcji aktywacji, a sieci splotowe często wykorzystują funkcję aktywacji ReLU, która nie jest ograniczona. Zresztą na rysunku zaprezentowanym na slajdzie 28 ewidentnie to widać. Jak doktorant chce rozwiązać ten problem? Ostatni aspekt dotyczy wyboru struktury sieci neuronowej. Dlaczego spośród wielu dostępnych rozwiązań (U-net, Inception, Yolo, itp.) doktorant zdecydował się na użycie akurat AlexNet? Poza tym, doktorant nie zaprezentował badań dotyczących wyboru hiperparametrów opisujących poszczególne warstwy splotowe, wybór metody grupowania (ang. *pooling*), dobór parametrów metody porzucenia (ang. *dropout*). Z całą pewnością zadanie nie zostało w pełni wykonane. Doktorant szacuje wykonanie zadania na poziomie 75% co należy uznać za mocno zawyżoną wartość.

Podsumowanie

Analizując dostarczoną dokumentację można wyciągnąć wniosek, że nie wszystkie zaplanowane zadania badawcze zostały w pełni wykonane. W szczególności razi brak wyników dotyczących realizacji zadania 3, którego głównym celem było zapoznanie się z przybornikiem Deep Learning środowiska obliczeniowego Matlab. Z drugiej strony, na slajdzie 28 doktorant zaprezentował strukturę sieci AlexNet zaprojektowaną z wykorzystaniem narzędzia Deep Network Designer, co świadczy o wykonaniu przez niego działań związanych z poznaniem przybornika Deep Learning. Jest również wiele do zrobienia w ramach dokończenia działań w ramach zadania 5. Szczęśliwie brakujące wątki z zadania 5 są ściśle powiązane z zadaniem 6, więc jest możliwość realizacji brakujących elementów zadania 5 wykonując zadanie 6.

Poza realizacją zadań badawczych, należy wskazać na aktywność doktoranta na polu wykraczającym poza Indywidualny Plan Badawczy. Doktorant pełni funkcję Przewodniczącego Samorządu Doktorantów Politechniki Opolskiej w roku akademickim 2021/2022. Jest również członkiem zespołu pracującego nad czasopismem naukowym Politechniki Opolskiej.

Podsumowując, pomimo pewnych niedociągnięć, które można w krótkim okresie czasu wyeliminować, Komisja pozytywnie ocenia zarówno rozwój naukowy doktoranta jak również jego postępy w realizacji Indywidualnego Planu Badawczego.

Członkowie komisji	Podpisy członków komisji:
1. Dr hab. inż. Stefan Wolny	…….
2. Porf. dr hab. inż. Krzysztof Patan	……
3. prof. dr hab. inż. Jerzy Skubis	…….