

I. Opis projektu

Wstęp

Rynek transportowy regulowany jest sporą ilością przepisów powiązanych z wieloma dziedzinami. Począwszy od przepisów ruchu drogowego, przez prawo transportowe, na prawie pracy skończywszy. Zastosowanie tak wielu aktów prawnych jest widoczne szczególnie w kontekście rozliczania czasu pracy kierowców w ruchu międzynarodowym. Proces ten wymaga jednoczesnego posługiwania się prawodawstwem unijnym oraz przepisami lokalnymi (dot. kraju, na terenie którego świadczona była praca). Dodatkowo, każda z firm spedycyjnych posiada własne regulaminy i zasady rozliczania czasu pracy. Specjaliści odpowiedzialni za realizację rozliczeń muszą więc biegle stosować kilkadziesiąt różnych aktów prawnych o różnej randze, z różnych krajów. Firma Legal Telematics jest podmiotem wyspecjalizowanym w prowadzeniu rozliczeń czasu pracy kierowców, poruszających się na trasach międzynarodowych.

Dziedzina

Źródłem danych wsadowych służących do prowadzenia rozliczeń są tachografy i systemy GPS oraz systemy kadrowe. Korelacja danych z nich pochodzących stanowi bazę dla dalszych działań, polegających na odpowiednim stosowaniu przepisów określających modele pracy, stawki, wykluczenia, przekroczenia, i wiele innych elementów, co ostatecznie daje końcowy wynik prowadzonych rozliczeń.

W ramach projektu B+R mają powstać innowacyjne narzędzia dedykowane obsłudze branży transportowej. Ich działanie włączone będzie w proces, jaki obecnie wykonują „ręcznie” specjaliści w kontekście analiz i przetwarzania danych z zastosowaniem w/w aktów prawnych.

Idea powstania nowego narzędzia wynika z palącej potrzeby rynku, ale przede wszystkim z możliwości jakie stwarzają najnowsze technologie, w tym struktury samouczące. Z ich pomocą mają zostać przezwyciężone największe problemy dotyczące podmiotów rozliczających czas pracy kierowców oraz dające zupełnie nowe możliwości, wynikające z szerszej niż dotąd integracji rozwiązań sprzętowych montowanych w pojazdach.

II. Obszar badawczy I: Zastosowanie AI w monitoringu zmian otoczenia prawnego procesu rozliczania czasu pracy kierowców w ruchu międzynarodowym.

Cel

Głównym celem zadania jest opracowanie narzędzi (w języku polskim i angielskim) z zakresu przetwarzania języka naturalnego, służących do analizy zapisów opublikowanych aktów prawnych lub propozycji ich zmian, dotyczących transportu samochodowego i pracy kierowców.

Zadaniem ww. narzędzi, opartych na wybranym/opracowanym Modelu AI jest nieustanne monitorowanie witryn internetowych w celu wyszukania informacji o zmianach (planowanych zmianach) lub utworzeniu nowych aktów prawnych w ww. zakresie.

Dopuszcza się oparcie ww. narzędzi na gotowych modelach AI, np. na modelu platformy AQUILA (<https://aquila.solutions/#model>) lub równoważnym, dedykowanym do zautomatyzowanej analizy dokumentów i aktów prawnych.

Warunkiem zastosowania modelu platformy AQUILA lub równoważnego będzie wykonanie dostosowania (adaptacji) jej możliwości do wymagań etapu/zadania. Dostosowanie to dotyczyć będzie m.in. uzupełnienia i dostosowania jej bazy wiedzy oraz ponownego trenowania jej modelu AI, tak aby spełniała oczekiwane funkcjonalności. Z uwagi na fakt, że dostępne platformy i ich modele przetwarzania języka naturalnego, bazują na określonych silnikach AI [np: IBM Watson Studio], trenowanie platformy również powinno bazować na usługach powiązanych z danym silnikiem, obejmujących m.in. takie techniki głębokiego uczenia jak uczenie transferowe (transfer learning).

Opracowanie odpowiednich technik uczenia oraz ich zaimplementowanie przy wykorzystaniu wybranego silnika będzie stanowiło główny problem badawczy do rozwiązania w zadaniu.

Dodatkowym problemem badawczym, do rozwiązania, będzie również zapewnienie i opracowanie odpowiednich danych źródłowych na potrzeby przygotowania ww. bazy uczącej. W tym celu analizie poddane powinny zostać portale publikacji odpowiednich aktów prawnych (lub ich propozycji),

obejmujących Prawo Polskie oraz Prawo UE oraz sposoby optymalnej ekstrakcji niezbędnych danych. W szczególności analiza ta dotyczyć będzie platform publikacji takich aktów prawnych jak:

1. Prawo Polskie

1.1. Kodeks pracy.

1.2. Kodeks karny.

1.3. Ustawa o czasie pracy kierowców.

1.4. Ustawa o wynagrodzeniu minimalnym.

1.5. Ustawa o delegowaniu pracowników w ramach świadczenia usług.

1.6. Ustawa o Państwowej Inspekcji Pracy.

1.7. Ustawa o drogach publicznych.

1.8. Ustawa o tachografach.

1.9. Ustawa o transporcie drogowym.

1.10. Rozporządzenie w sprawie należności z tytułu podróży służbowej.

1.11. Minimalne wynagrodzenie za pracę oraz wysokości minimalnej stawki godzinowej w 2020 r.

1.12. Minimalne wynagrodzenie za pracę oraz wysokości minimalnej stawki godzinowej w 2021 r.

1.13. Minimalne wynagrodzenie za pracę oraz wysokości minimalnej stawki godzinowej w 2022 r.

2. Prawo UE.

2.1. Dyrektywa (UE) 2020/1057 dotycząca delegowania pracowników w transporcie oraz kontroli.

2.2. Rozporządzenie (UE) 2020/1055 wprowadzające zmiany w dostępie do rynku (1071/2009 oraz 1072/2009).

2.3. Rozporządzenie (UE) 2020/1054 wprowadzające zmiany w czasie jazdy i odpoczynku (561/2006) oraz tachografach (165/2014).

2.4. Rozporządzenie Komisji (UE) nr 2016/403 zawierające klasyfikacje naruszeń.

2.5. Dyrektywa 96/71/WE dotycząca delegowania pracowników w ramach świadczenia usług (2020).

2.6. Rozporządzenie (WE) nr 1073/2009.

2.7. Rozporządzenie (WE) nr 1072/2009 (2022) dotyczące wspólnych zasad dostępu do rynku międzynarodowych przewozów drogowych.

2.8. Rozporządzenie (WE) nr 1071/2009 (2022) dotyczące warunków wykonywania zawodów przewoźnika drogowego.

2.9. Dyrektywa 2002/15/WE.

2.10. Umowa AETR.

2.11. Rozporządzenie Komisji (UE) nr 581/2010 dot. częstotliwości pobierania danych.

2.12. Rozporządzenie Rady (EWG) 3821/85.

2.13. Rozporządzenie (UE) nr 165/2014 (2020).

2.14. Rozporządzenie (WE) nr 561/2006 (2020).

Celem powyższej analizy ma być będzie opracowanie tzw. Crawlera internetowego, który zapewnić będzie stałe monitorowanie i odwzorowywania aktualnego stanu prawnego w ww. zakresie. Oczekiwanym elementem tego opracowania będzie również odpowiednia struktura bazy danych, w której dynamicznie odwzorowywany będzie ww. stan prawny. Opracowanie tych narzędzi stanowi kolejny problem badawczy do rozwiązania w zadaniu.

Crawler internetowy oraz model AI, o którym mowa powyżej, mają być podstawą organizacji tzw. Modułu monitorowania aktów prawnych, którego zadaniem, w szczególności będzie:

1. stałe monitorowanie oraz weryfikacja źródeł informacji prawnej, z podziałem na:

1.1. oficjalne źródła, których obowiązkiem jest obwieszczenie aktów prawnych.

1.2. oficjalne źródła, które nie posiadają kompetencji do obwieszczenia aktów prawnych.

1.3. nieoficjalne źródła, takie jak portale branżowe, grupy dyskusyjne, artykuły naukowe.

2. przygotowanie opisu wyszukanych informacji wraz z rekomendacjami konkretnych działań, jakie należy podjąć aby zapewnić ciągłość pracy systemu w zgodzie z porządkiem prawnym.

Uzupełnieniem dla ww. modułu ma być Moduł Notyfikacji Użytkownika, którego funkcjonalności obejmować będą:

1. powiadamianie użytkowników systemu lub zdefiniowane kontakty o wystąpieniu określonych faktów w ramach działania modelu monitorowania aktów prawnych,

2. definiowalny zestaw reguł pozwalających na określenie warunków, które muszą zajść w ramach monitorowanych faktów, aby wyzwolone zostało wysłanie powiadomienia do odbiorcy,

3. powiadamianie o zmianach w ramach systemu, a także za pośrednictwem wiadomości e-mail.

III. Obszar badawczy II: Zastosowanie AI w analizach danych i modelowaniu behawioralnym.

Dane pozyskiwane z modułów GPS do rozliczania czasu pracy kierowców stanowią niewielki odsetek danych jakie gromadzone są w systemach pokładowych pojazdów. Opracowanie nowego profilu sprzętowego oraz umożliwienie jego głębszej integracji z systemami dedykowanymi analizom danych jest podstawą do pełniejszego wykorzystania danych z pokładu pojazdów. Przedmiotem prac B+R ma być:

1. Opracowanie rozwiązania umożliwiającego komunikację 2stronną z częścią sprzętową, przetwarzającego dane pochodzące z pojazdu i umożliwiającą komunikację zwrotną z kierowcą lub operatorem systemu.
2. Tworzenie profili behawioralnych poszczególnych kierowców, na podstawie danych pozyskiwanych z szyny CAN (za pomocą modułu GPS oraz za pomocą procedury warsztatowej), w formie pakietów danych obejmujących określone przedziały czasu. Profile behawioralne tworzone mają być w obszarach:
 1. Bezpieczeństwa w ruchu drogowym.
 2. Wydajności pracy.
 3. Ecodrivingu.

Cel

W ramach zadania przewiduje się realizację następujących funkcjonalności:

1. monitoring floty pojazdów pozwalający na stworzenie dynamicznej mapy, na której będzie można sprawdzić trasy kierowców oraz odstępstwa od wcześniej planowanych tras. Jest to niezbędne do analiz ryzyk związanych z dopuszczeniem, na określonych odcinkach czy elementach infrastruktury drogowej, określonych ciężarów czy gabarytów. Monitoring ten także ma być podstawą określania poziomu przestrzegania bezpiecznych prędkości na danym odcinku. Wykrywać też należy sytuacje związane z gwałtownymi akcjami (np. hamowaniem lub przyspieszaniem)
2. analiza danych historycznych pozwoli na ocenę postępów poszczególnych pracowników dotyczących ich stylu jazdy, a także sprawdzenie jakie trasy są przez nich wybierane oraz jak bardzo ekonomicznie prowadzą oni swój pojazd. Wysoki wynik będzie świadczył o znajomości przepisów, zachowawczym i ekonomicznym stylu jazdy, a także umiejętności przewidywania trudnych sytuacji na drodze.
3. detekcja niebezpiecznych zachowań ma umożliwić wykrycie groźnych manewrów, takich jak gwałtowne skręty i hamowania oraz przyspieszanie, szczególnie w korelacji z panującymi warunkami pogodowymi - np. podczas opadów deszczu lub śniegu
4. tworzenie raportów, w tym rankingu kierowców, na podstawie danych gromadzonych inkrementalnie w bazie danych. Raporty powinny być definiowalne przez operatora systemu i obejmować wszystkie gromadzone dane.
5. stworzenie API dla aplikacji mobilnej stanowiącej kanał zwrotny dla kierowców. API ma umożliwiać wysyłanie raportów, komunikatów, notyfikacji, stanowiących element komunikacji systemu z kierowcami.

W celu rozpoznania odpowiednich wzorców związanych ze stylem jazdy konieczne będzie zebranie danych z wielu godzin jazdy. Część z tych danych będzie musiała być otagowana w celu oznaczenia z jakim zdarzeniem lub stanem wiąże się dana obserwacja. Dla ułatwienia tego procesu należy skorzystać z techniki klasteryzacji (np. k-means) dla wstępnego podziału danych na podzbiory. W innych przypadkach dane z czujników będą powiązane z danymi ciągłymi, np. związanymi z trajektorią pojazdu. Przykładowymi parametrami jakie będą poddawane analizie są: prędkość pojazdu, pozycja pedałów gazu i hamulca, kąt obrotu kierownicy, czujnik deszczu, a także dane związane z trajektorią pojazdu (GPS). Wśród metod, które będą mogły zostać zastosowane należy wyróżnić ukryty model Markowa (ang. hidden Markov model) do reprezentowania zależności pomiędzy obserwacjami stanów ciągłych i dyskretnych. Do określania zmęczenia lub senności kierowcy, a także klasyfikacji jego zachowania przewiduje się wykorzystanie maszyny wektorów nośnych (ang. Support Vector Machine - SVM), sieci Bayesowskie (ang. Bayesian networks), drzewa decyzyjne (ang. decision trees) czy logikę rozmytą (ang. fuzzy logic). Należy również wykorzystać algorytmy głębokiego uczenia (ang. deep learning), które będą mogły zostać użyte do przewidywania zachowania kierowcy w oparciu o dane telematyczne pojazdu. Innym zastosowaniem będą konwolucyjne sieci neuronowe (ang. Convolutional Neural Networks - CNNs), które będą mogły wspomóc ocenę stanu zmęczenia (senności) kierowcy lub określenie aktualnych warunków pogodowych.

IV. Wymagania technologiczne

Poszczególne usługi realizowane w ramach prac B+R mogą być implementowane różnych technologiach np. .NET, Python lub równoważnych.

W przypadku modułu analitycznego realizującego różnorodne zadania, np:

- wyszukiwanie danych dotyczących konkretnego kierowcy i generowanie odpowiedniego raportu,
- uwzględnienie stylu jazdy kierowcy wpływającego na bezpieczeństwo i ekonomikę jazdy.
- automatyczne wyszukiwanie zmian i śledzenie relacji pomiędzy regulacjami prawnymi w języku polskim i angielskim dotyczących transportu samochodowego i pracy kierowców oraz powiadamianie o odnośnych zmianach.

oczekuje się możliwości zaimplementowania rozwiązań w środowisku chmurowym (bez tzw. vendorlockingu w chmurze publicznej lub prywatnej) z zapewnieniem wymaganego poziomu cybersecurity, adekwatnego dla danych osobowych i biznesowych.

Poza typowym przetwarzaniem danych hurtowych, zostanie zastosowana technologia AI np.: algorytmy uczenia maszynowego. Przykładem takich algorytmów są np: modele ARIMA (AutoRegressive Integrated Moving Average) lub LSTM (Long Short-Term Memory), do analizy danych historycznych, których podstawową miarą jest czas.

Przy pracach nad modelami ML przewidywana jest również ewentualna augmentacja danych służących do budowania ciągów uczących. W toku prac badawczych może bowiem okazać się konieczne generowanie danych uczących które zoptymalizują finalne modele.

Środowiskiem planowanym do budowania oraz korzystania z modeli będzie platforma Python.

Ponieważ w praktyce dla .NET, nie istnieje natywna biblioteka do ARIMA i LSTM w razie stosowania tego typu rozwiązań dopuszcza się stosowanie bibliotek, które umożliwiają integrację z językiem Python w aplikacjach .NET np: IronPython stanowiąca implementację języka Python działająca na platformie .NET.

Proponowane narzędzia w planowanej architekturze rozwiązania (wskazane poniżej lub równoważne)

Software

- Usługa WEB-API (RestFul)
- Platforma .NET core/python
- Baza danych PostgreSQL możliwość replikacji i skalowania
- Obiektowy magazyn plików

Hardware

- Środowisko wirtualizacji zasobów CPU/GPU