AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE
Wydział Elektrotechniki, automatyki, informatyki i inżynierii biomedycznej.

Praca dyplomowa magisterska

Federacyjne uczenie maszynowe

Federated machine learning

Autor: *Mikołaj Skrzyniarz*
Kierunek studiów: *Informatyka i Systemy Inteligentne*
Opiekun pracy: *dr inż. Piotr Szwed*

Kraków, 2022

Spis treści

1. Wstęp

Dziedziny zajmujące się zagadnieniami związanymi z sztuczną inteligencją istnieją w świecie technologii i nauki od dłuższego czasu. Ich popularność stale rośnie a sztuczna inteligencja powiązana jest z takimi obszarami jak matematyka, statystyka, nauki o danych (ang. *data science)*, duże zbiory danych (ang. *big data)* oraz z oczywistych względów – informatyka (ang. *computer science*). Trudno się temu dziwić, wszystkie te zagadnienia niejako przeplatają się nawzajem – poruszają podobne problemy oraz wymagają zbliżonych umiejętności.

W przeciągu ostatnich dwudziestu lat termin sztuczna inteligencja ewoluował z nauki wzbudzającej ogromną ciekawość wśród ludzi do praktycznej technologii, powszechnie używanej w celach komercyjnych. Ten postęp spowodowany jest rosnącą dostępnością systemów umożliwiające prowadzenie różnego rodzaju skomplikowanych obliczeń, materiałów naukowych oraz świadomością jak i wiedzą uczonych, studentów jak i wszystkich pracowników branż technologicznych. Ta z kolei prowadzi do powstawania nowych, bardziej rozbudowanych i zaawansowanych algorytmów. W efekcie liczba firm stosujących rozwiązania bazujące na mechanizmach sztucznej inteligencji w ubiegłych czasach wzrosła diametralnie. Wizja komputerowa (ang. *computer vision)*, przetwarzanie języka naturalnego, rozpoznawanie mowy, wykrywanie schorzeń oraz różnego rodzaju anomalii to tylko niektóre z wielu przykładów użycia, a z prostych mechanizmów wspomagających podejmowanie decyzji korzystamy, świadomie lub nie, tak naprawdę na co dzień.

Uczenie maszynowe jest największym obszarem szeroko rozumianej sztucznej inteligencji. Może być rozumiany jako zajmujący się szukaniem rozwiązania problemu polegającego na stworzeniu urządzenia, które będzie osiągało lepsze wyniki w skali ustalonej wcześniej metryki poprzez naukę w oparciu o doświadczenie oraz decyzje podejmowane w przeszłości. Przykładem może być wykrywanie schorzenia na podstawie zdjęć rentgenowskich danego narządu. Celem jest poprawne przydzielenie etykiet „zdrowy" i „chory" dla każdego ze zdjęć. By to osiągnąć dany algorytm musi nauczyć się rozróżniać zdjęcia na podstawie określonych cech. To z kolei odbywa się w procesie uczenia, podczas którego dany model poddaje się próbom na zbiorze uczącym, zawierającym dane służące jako wzorzec, mające już przydzielone etykiety [2].

Tradycyjny proces uczenia z uwzględnieniem danych zbieranych urządzenia rozproszone polega na agregacji takich danych na wspólnym urządzeniu pełniącym rolę serwera, wytrenowaniu modelu oraz propagacji gotowego modelu między urządzeniami. Głównym problemem takiego modelu jest fakt, że dane istnieją w formie odizolowanych obiektów oraz ich bezpieczeństwo jak i prywatność, które mogą być naruszone podczas procesu agregacji. Rozwiązaniem tego może być zastosowanie federacyjnego uczenia maszynowego, zaproponowanego przez firmę kilka lat temu przez firmę *Google* [3]. W teorii skuteczność wytrenowanego w ten sposób modelu

powinna być zbliżona do tradycyjnej metody przy zachowaniu większego bezpieczeństwa całego procesu oraz niższego ryzyka wycieku danych.

1.1. Cel pracy

Celem niniejszej pracy było zaprojektowanie oraz przeprowadzenie eksperymentów mających na celu symulację federacyjnego uczenia maszynowego. Wyniki powinny zostać porównane z wynikami uzyskanymi przy użyciu tradycyjnego sposobu uczenia. W wykorzystanym podejściu samodzielni agenci trenują swoje modele przy użyciu obserwowanych danych. Modele te są okresowo agregowane, a następnie dystrybuowane wewnątrz grupy.

W pracy należało rozważyć różne tryby działania algorytmu, a także zaproponować metodę agregacji wag oraz parametry sterujące. Docelowy system miał być przetestowany na dużym zbiorze danych z użyciem modeli o różnym stopniu złożoności. W przypadku zastosowania *transfer learning* rdzeń modelu nie powinien podlegać modyfikacji.

1.2. Zakres pracy

TODO

2. Wprowadzenie teoretyczne

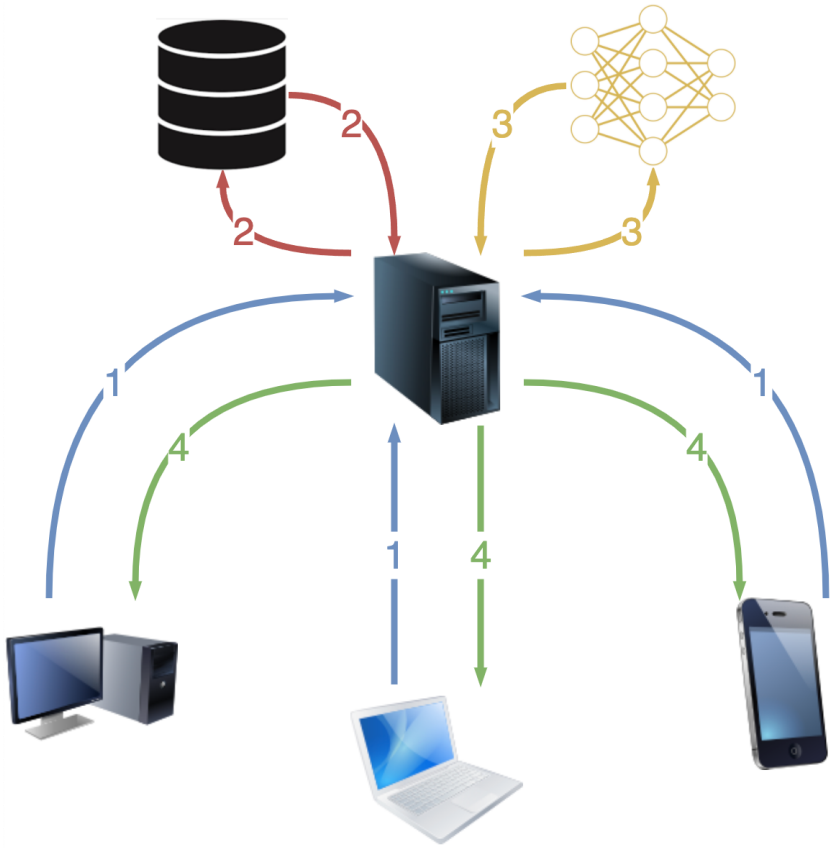
Celem drugiego rozdziału pracy jest zaznajomienie czytelnika z teorią na temat poruszanych pojęć oraz zagadnień. W oparciu o wykorzystane pozycje literaturowe wytłumaczone zostanie czym tak naprawdę jest federacyjne uczenie maszynowe, jakie są różnice względem tradycyjnego modelu uczenia oraz jakie wyróżniamy rodzaje opisywanej metody.

2.1. Rozproszone uczenie maszynowe

Tradycyjny model uczenia maszynowego nazywany jest także rozproszonym uczeniem maszynowym (ang. *distributed machine learning*). Jest to to wielowęzłowy system gdzie każdy węzeł symbolizuje pojedyncze urządzenie. Taka struktura umożliwia przetwarzanie dużych zbiorów danych, których powiększanie odbywa się poprzez dokładanie kolejnych węzłów. Taki sposób uczenia przewiduje agregację danych z wielu urządzeń w jeden zbiór, który następnie służy do nauki modelu określonego typu. Tak skonstruowany model zostaje następnie wysłany do każdego urządzenia korzystającego z danej aplikacji. Samo urządzenie zbiera dane w czasie rzeczywistym, a proces ich wysyłania do jednostki na której ma miejsce proces uczenia odbywa się okresowo. Jest to jedną z wad tego rodzaju uczenia – nie ma możliwości by zbierane przez urządzenie dane na bieżąco aktualizowały używany model. Niemniej jednak cały proces przeprowadzony w opisywany sposób skutkuje rosnącą skutecznością danego algorytmu. Taki typ uczenia jest w dobry pod względem uzyskiwanej skuteczności jednak wzbudza pewne obawy co do prywatności danych. Wszystkie urządzenia mobilne, IoT (ang. *internet of things*) kolekcjonują całą masę danych, w związku z czym zachowanie ich prywatności staje się co raz większym wyzwaniem.

Proces uczenia maszynowego na podstawie danych zebranych przez wiele urządzeń rozproszonych w oparciu o tradycyjną metodę możemy w uproszczeniu podzielić na cztery kroki (patrz rys. 2.1):

1. okresowe wysyłanie zgromadzonych przez urządzenia rozproszone danych do jednostki centralnej;
2. agregacja uzyskanych danych w jeden zbiór uczący;
3. przeprowadzenie procesu uczenia na serwerze;
4. dystrybucja zbudowanego modelu sztucznej inteligencji między urządzeniami korzystającymi z systemu [2].



Rys. 2.1. Diagram przedstawiający model tradycyjnego uczenia maszynowego.

2.2. Federacyjne uczenie maszynowe

Koncept federacyjnego uczenia maszynowe pierwszy raz światło dzienne ujrzał w 2016 roku, kiedy został zaprezentowany przez firmę *Google*. Ten nowy na tamte czasy model uczenia został użyty w aplikacji *Google keyboard* gdzie w sposób kolaboracyjny trenowano model na podstawie kilku urządzeń z systemem *Android*. Jednak federacyjny sposób uczenia może zostać zaimplementowany użyciu każdego z urządzeń używanych w obszarze sztucznej inteligencji. Sama idea ma duży potencjał zrewolucjonizować rynek *AI*. Ciekawym przykładem może być sytuacja, gdzie technolodzy oraz pracownicy branży medycznej z całego świata w sposób federacyjny trenowali model mający na celu wykrywać chorobę COVID-19 na podstawie skanów klatki piersiowej [4]. Głównym celem dla którego podjęto próby wynalezienia nowego sposobu uczenia maszynowego było zapewnienie bezpieczeństwa oraz zniwelowanie ryzyka utraty lub wycieku danych, używając zestawów danych znajdujących się na wielu urządzeniach lub w wielu organizacjach.

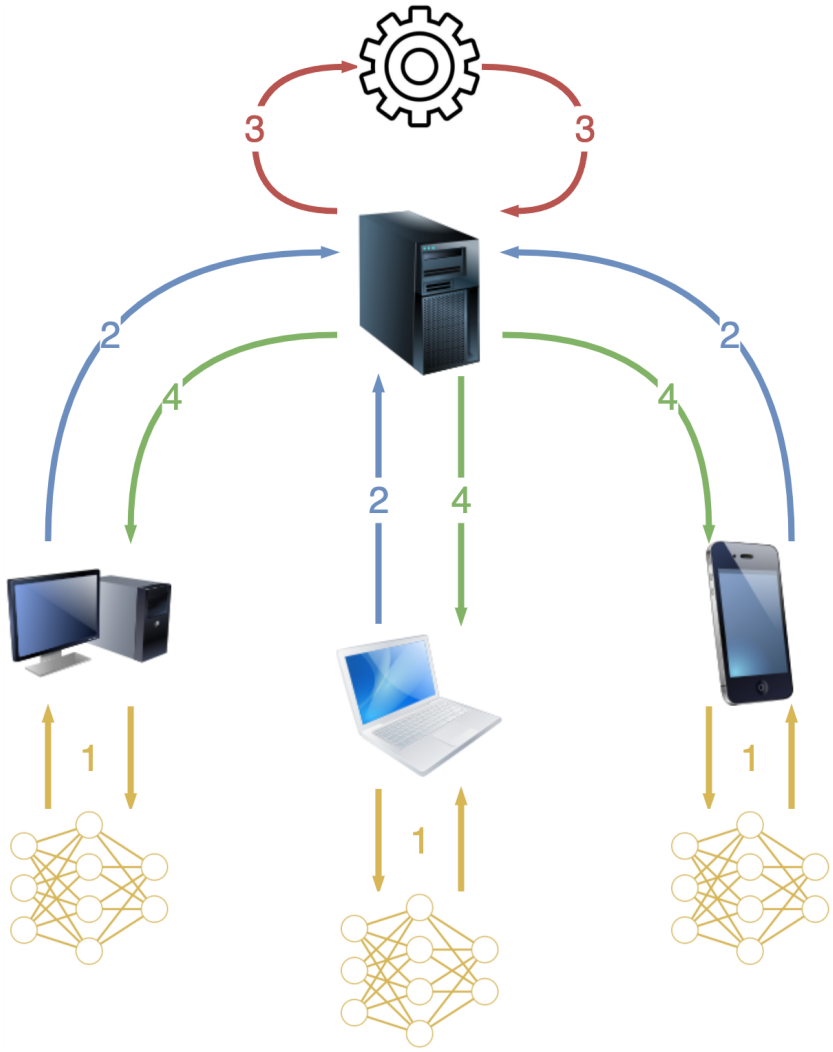
Taki sposób tworzenia modeli sztucznej inteligencji, w przeciwieństwie do wspomnianego wcześniej tradycyjnego, nie wymaga by zgromadzone przez urządzenia dane opuszczały pamięć danego urządzenia. Zamiast tego szkolenie przebiega lokalnie, przy użyciu zebranych danych. Odbywa się to, tak jak w przypadku modelu

tradycyjnego, cyklicznie, jednak ze względu na brak konieczności wysyłania danych do serwera interwały są w większości przypadków dużo krótsze, a sam czas szkolenia jest mniejszy ze względu na mniejsze przyrosty zbioru danych. Wytrenowane w ten sposób modele są następnie wysyłane do jednostki centralnej, gdzie poddawane są agregacji. W efekcie czego z wielu takich modeli uzyskiwany jest jeden, wyjściowy model, który jest dystrybuowany pomiędzy urządzeniami. Przy zastosowaniu konfiguracji w której kilka organizacji wspólnie pracuje nad modelem sztucznej inteligencji sytuacja wygląda w sposób analogiczny – parametry modeli są agregowane na współdzielonym serwerze lub chmurze. Następnie model wynikowy jest wysyłany do organizacji, które dokonują aktualizacji lokalnie.

Docelowo przy zastosowaniu takiej metody zbiór danych znajdujący się na danym urządzeniu nie powinien być w żaden sposób udostępniony pozostałym urządzeniom. Niemniej dopuszczalne są pewne odstępstwa od tej normy, przy czym muszę w takiej sytuacji zostać ustalone i zachowane odpowiednie procedury bezpieczeństwa. Skuteczność uzyskanego w sposób federacyjny algorytmu powinna być zbliżona do algorytmu stworzonego na tym samym zbiorze danych przy użyciu tradycyjnego sposobu uczenia.

Analogicznie do tradycyjnego modelu uczenia maszynowego, proces federacyjnego uczenia możemy w uproszczeniu podzielić na cztery etapy (patrz rys. 2.2):

1. cykliczne przeprowadzanie nauki modelu na podstawie gromadzonych lokalnie danych;
2. okresowe wysyłanie wyszkolonych modeli do serwera;
3. agregacja uzyskanych modeli w jeden, wynikowy model;
4. dystrybucja zbudowanego modelu sztucznej inteligencji między urządzeniami korzystającymi z systemu [3].



Rys. 2.2. Diagram przedstawiający model federacyjnego uczenia maszynowego.

2.3. Prywatność danych w uczeniu federacyjnym

Prywatność i bezpieczeństwo danych są kluczowymi właściwościami federacyjnego uczenia maszynowego. Ich zapewnienie wymaga istnienia różnego rodzaju modelów bezpieczeństwa i analizy. Poniżej wyróżniono kilka przykładowych modeli.

Prywatność różnicowa (ang. *Differential Privacy)* polega na dodaniu szumu do danych lub użyciu metod generalizacji co ma na celu ukrycia pewnych wrażliwych cech danego zbioru. W wyniku takiego działania pojedyncze próbki mogą być trudne do rozróżnienia, a sam algorytm uczenia działa na danych przybliżonych do rzeczywistych.

Szyfrowanie homomorficzne (ang. *Homomorphic Encryption)* jest to typ algorytmów szyfrujących, pozwalających na przeprowadzanie obliczeń przy użyciu zaszyfrowanych danych, bez konieczności ich deszyfrowania. Ze względu na fakt, że klucz deszyfrujący znany jest tylko przez urządzenie źródłowe, ryzyko wycieku danych jest wyjątkowo niskie. Powszechnie stosowane w algorytmach uczenia maszynowego.

Zastosowanie tego modelu w większości przypadków skutkuje uzyskaniem kompromisu pomiędzy dokładnością a bezpieczeństwem i prywatnością danych.

Bezpieczne obliczenia wielopartyjne (ang. *Secure Multi-party Computation, SMC)*. Tego typu modele w sposób naturalny angażują wiele obiektów, przy czym głównym założeniem jest, że każdy z obiektów zna tylko swoje dane wejściowe oraz wyjściowe. Zerowa wiedza jest szeroko pożądana w przypadku federacyjnego uczenia maszynowego, jednak często wymaga ona skomplikowanych protokołów obliczeniowych w związku z czym osiągniecie takiego stanu nie jest proste. Niemniej jednak w określonych warunkach częściowa wymiana wiedzy między danymi instancjami może być dopuszczalna pod warunkiem, że są zachowane odpowiednie działania mające na celu zachowanie bezpieczeństwa. Takie rozwiązanie w niektórych przypadkach może okazać się optymalnym kompromisem pomiędzy zapewnieniem wystarczającego bezpieczeństwa oraz wydajności danego systemu [3].

2.4. Typy federacyjnego uczenia

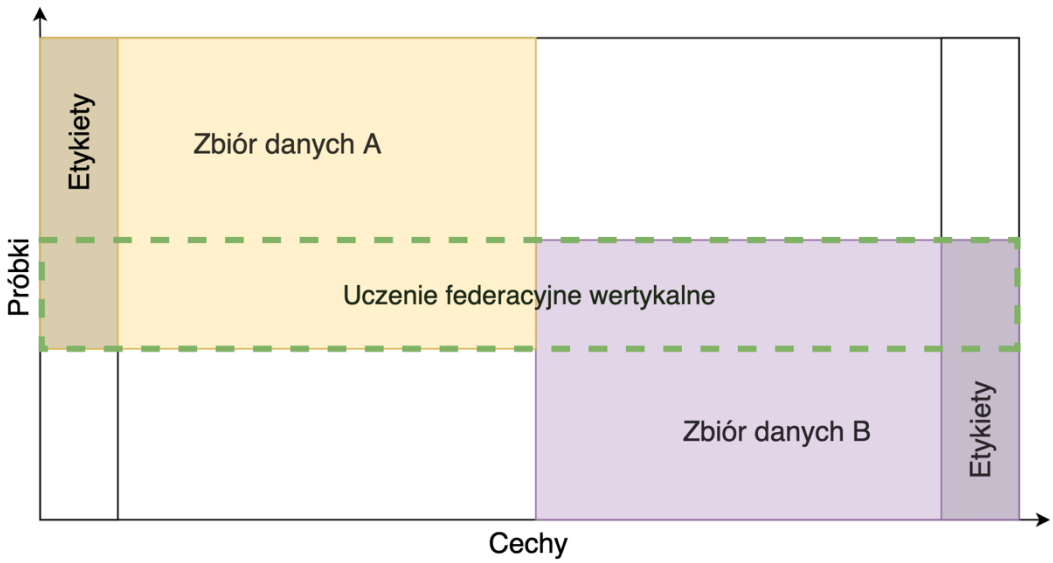
Aktualnie jesteśmy w stanie zaobserwować dwa sposoby kategoryzacji typów uczenia federacyjnego:

- ze względu na wspólne aspekty udostępnianych przez urządzenia zbiorów danych,
- ze względu na charakter decentralizacji źródeł danych.

Aktualnie nie wszystkie z opisanych rozwiązań są powszechnie używane w środowiskach produkcyjnych jednak widoczna jest tendencja wzrostowa. Poniżej opisane zostały typy należące do pierwszej z kategorii.

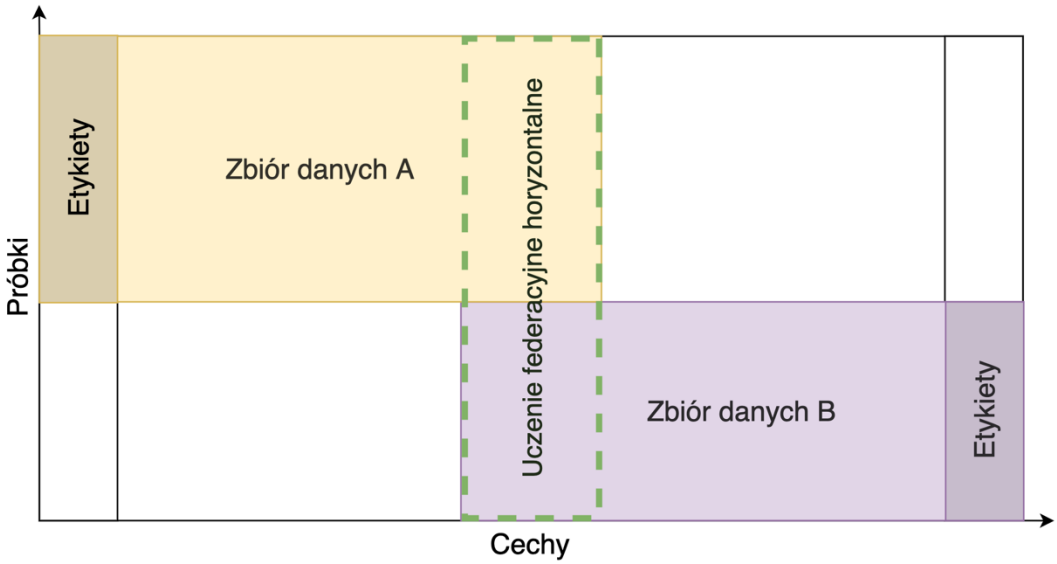
Uczenie federacyjne wertykalne (ang. *Vertical Federated Learning)* polega na trenowaniu modeli na podstawie danych pochodzących z różnych źródeł, opisujących różne dane kontekstowe dotyczące tych samych obiektów. Jako przykład posłużyć mogą przychodnia lekarska i lokalny bank znajdujące się w niewielkiej miejscowości. Najprawdopodobniej z obu tych instytucji korzysta spora część mieszkańców, zatem identyfikatory próbek, na przykład numer pesel, będą w obu przypadkach takie same. Jednak ze względu na różny charakter obu działalności badane atrybuty będą zupełnie inne. Wertykalny rodzaj uczenia maszynowego polega na agregacji parametrów modeli przeznaczonych do operowania na zbiorach danych opisujących różne cechy i zbudowaniu globalnego modelu, gdzie każda ze stron będzie brała czynny udział w procesie uczenia. Docelowo tak uzyskany model mógłby być z powodzeniem stosowany w każdej z tych organizacji. Ten rodzaj szkolenia zakłada, że agenci mogą zagrażać sobie nawzajem pod kątem bezpieczeństwa danych. Niemniej jednak każda ze stron jest niezależna ma ryzyka nawiązania współpracy przez jakiekolwiek z nich. Po zakończeniu procesu uczenia każda ze stron otrzymuje model posiadający parametry

związane z cechami, które opisywał wysłany zbiór danych w związku z czym każda ze stron musi brać udział w podczas wnioskowania (ang. *inference time*).



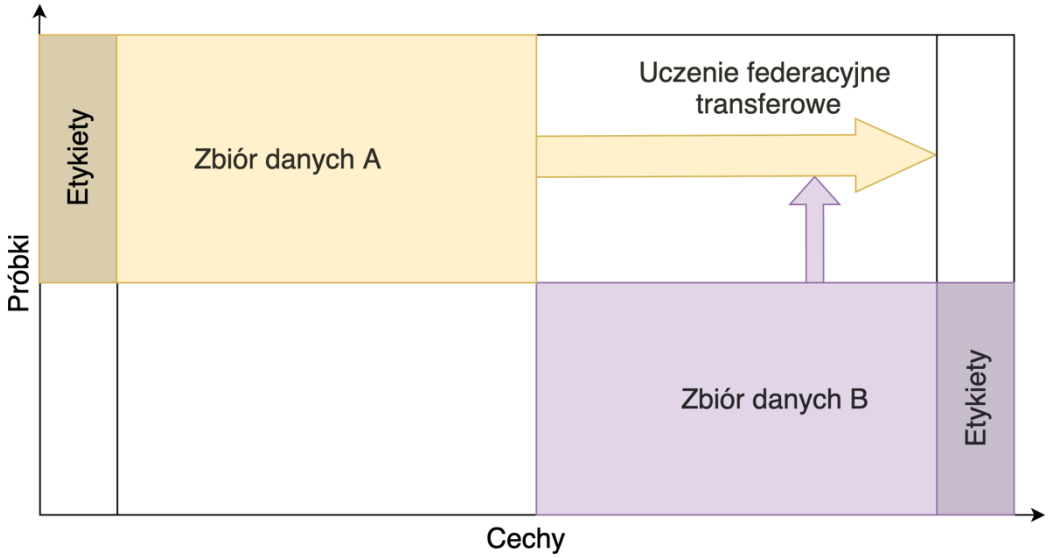
Rys. 2.3. Uczenie federacyjne wertykalne.

Uczenie federacyjne horyzontalne (ang. *Horizontal Federated Learning)* jest to proces uczenia w sytuacji gdy zbiory danych pochodzących z różnych źródeł opisują te same dane kontekstowe, dotyczące jednak różnych zdarzeń lub obiektów. Przykładem mogą być dwie prywatne przychodnie lekarskie znajdujące się w różnych miastach. Z dużym prawdopodobieństwem można założyć, że będą one zbierać te same lub bardzo podobne dane na temat swoich pacjentów. Jednak w związku z położeniem w dwóch różnych miejscowościach w skład pacjentów obu przychodni będą wchodziły inne osoby. Przykładem mogą być wszystkie organizacje działające w tej samej branży, mające jednak inne grupy odbiorców. Podczas zastosowania tego typu uczenia możliwa jest też wymiana danych na temat danej grupy cech. Jest to najczęściej używany rodzaj uczenia, stosowany powszechnie przez *Google*, na przykład w aplikacji *Google Assistant* [6]. Pozwala on na zachowanie bezpieczeństwa oraz prywatności między użytkownikami, a jedynym obiektem w całym systemie mogącym stwarzać zagrożenie jest serwer. Po zakończeniu procesu uczenia wszystkie parametry modelu wynikowego udostępniane są dla każdego odbiorcy.



Rys. 2.4. Uczenie federacyjne horyzontalne.

Uczenie federacyjne transferowe (ang. *Federated Transfer Learning*) może zostać zaimplementowany w sytuacji gdy dwie organizacje opisują zarówno różne cechy jak i obiekty. Przykładem mogą być bank oraz sklep z elektroniką znajdujące się w innych krajach. Ze względu na dużą odległość między obydwoma miejscami najprawdopodobniej grupa odbiorców będzie zupełnie inna. Ze względu na różne branże obu organizacji tylko bardzo mała część gromadzonych danych kontekstowych będzie miała taki sam charakter. Ten typ uczenia służy raczej jako rozszerzenie do wspomnianych wcześniej uczenia wertykalnego i horyzontalnego, aniżeli jako samodzielne rozwiązanie [3].



Rys. 2.5. Uczenie federacyjne transferowe.

Uczenie federacyjne z zastosowaniem wielu silosów (ang. *Cross-Silo Federated* Learning) oraz uczenie federacyjne z zastosowaniem wielu urządzeń (ang. *Cross-Device Federated Learning*) należą do drugiej kategorii według której dzielimy typy

uczenia federacyjnego. Pierwszy z nich zakłada istnienie „silosów" gromadzących dane takich jak organizacje rób regiony geograficzne. Przy takim ustawieniu liczba klientów wynosi waha się zazwyczaj od 2 do 100. Z kolei drugi zakłada istnienie od 0 do nawet 10^{10} klientów, jednak w tej sytuacji mała część z nich bierze czynny udział w procesie budowania modelu wyjściowego. Uczenie z zastosowaniem wielu silosów jest coraz częściej stosowane, na przykład w takich branżach jak medyczna lub finansowa [5].

Opis przebiegu wykonanych badan i eksperymentów

Podsumowanie

Bibliografia

Opracowania książkowe

[1] TODO

Dokumenty

[2] *Machine learning: Trends, perspectives, and prospects*, M.I. Jordan, T.M. Mitchell

[3] *Federated Machine Lraning: Concept and Applications,* Yang Liu, Tianjian Chen, Yongxing Tong

[4] *Federated learning: Opportunities and Challenges,* Priyanka Mary Mammen

[5] *Cross-silo federeated training in the cloud with diversity scalind and semi-supervised learning,* Kishore Nandury, Anand Mohan, Frederick Weber

Źródła internetowe

[6] Google hey google, https://support.google.com/assistant/answer/10176224?hl=en