

Celem projektu jest stworzenie modelu sztucznej inteligencji do przewidywania wstrząsów wtórnych na podstawie pomiarów seismologicznych poprzednich zjawisk sejsmicznych oraz porównanie go z innymi istniejącymi rozwiązaniami. Wykrycie pierwszych zdarzeń sejsmicznych jest zbyt nieprzewidywalnym zagadnieniem, aby próbować je przewidzieć, natomiast w przypadków wstrząsów wtórnych mamy dane na temat wstrząsów głównych. W celu eksploracji oraz porównania skuteczności różnych podejść, wykorzystano trzy rodzaje sieci neuronowych: MLP, CNN oraz LSTM. Ten system pozwoliłby na wsparcie służb ratunkowych w organizacji pomocy poszkodowanym i ostrzegałby przed wtórnymi wstrząsami. Z racji uszkodzeń budynków po wcześniejszych falach sejsmicznych, tego typu wstrząsy są równie niebezpieczne, szczególnie w przypadku w sytuacji kiedy są silniejsze od poprzednich.

Istniejące rozwiązania

Podstawową działania obecnie używanych systemów ostrzegawczych, nadal jest jak najwcześniejsze wykrycie zdarzenia sejsmicznego. Z tego powodu tereny, które znajdują się w Epicentrum nie dostaną żadnego ostrzeżenia. Dotyczy to również wstrząsów wtórnych. Systemy te wykrywają szybsze, ale mniej niszczycielskie fale podłużne "P", które pojawiają się przed falami poprzecznymi "S". Oprócz tych rozwiązań stworzono kilka modeli sztucznej inteligencji. Jedne próbują wspomóc już istniejące systemy wykrywania, a inne próbują je przewidzieć. W przeciwieństwie do omawianego projektu, skupiają się one na analizie metadanych, jak lokalizacja czy czas wystąpienia, a nie bezpośredni zapis sejsmografu. Dodatkowo, mają ograniczenia, jak działanie na małym obszarze, lub przewidują tylko ilość w danym okresie czasowym.

STEAD

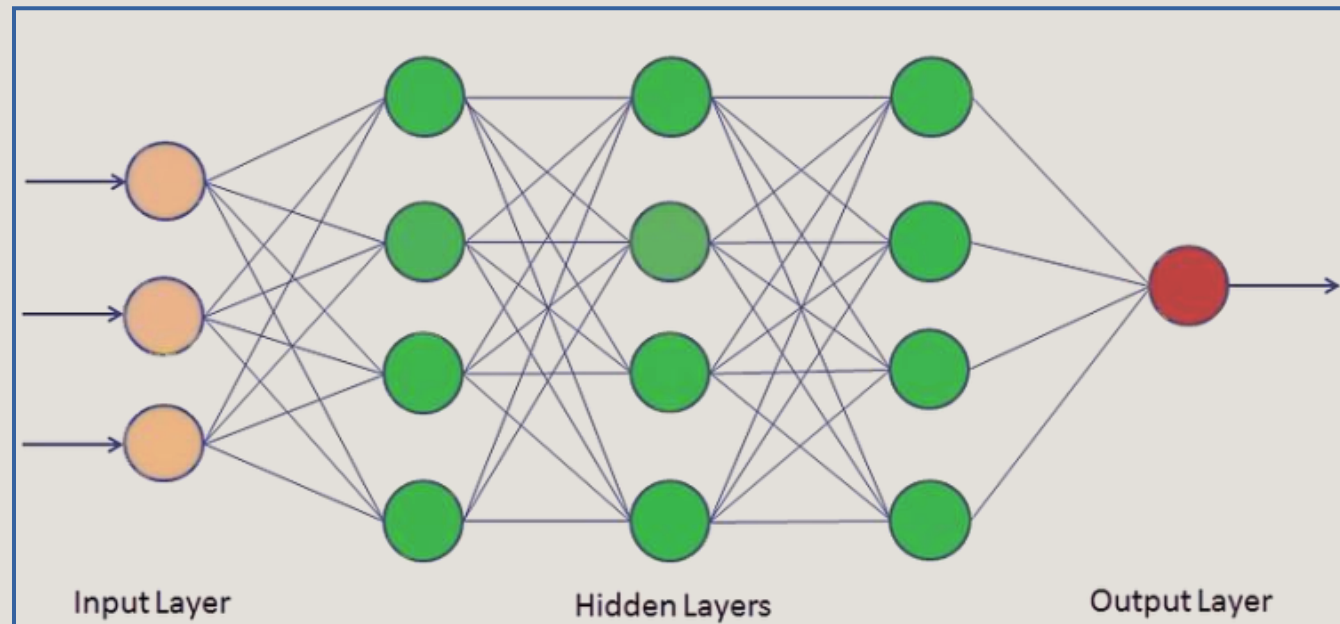
Przez długi czas prace nad tego typu projektami były powstrzymywane przez brak odpowiedniego zbioru danych. Do 2019 potrzebne informacje były rozproszone wśród naukowych instytucji, często państwowych, prowadzących badania geologiczne. Dopiero w roku 2019 został opublikowany STEAD [1], "Stanford earthquake dataset". Jest to zbiór danych o falach sejsmicznych z całego świata zmierzonych w latach od 1984 do 2018 roku. Każdy sejsmogram, jest opisany miejscem i czasem wystąpienia epicentrum, magnituda, oraz informacji o czujniku je rejestrującym. Zawiera on ponad miliona takich zapisów, oraz ponad stu tysięcy zapisów drgań niskiej amplitudy, tła, wynikających z naturalnych procesów czy działalności człowieka.

Powstałe modele

W ramach niniejszego projektu opracowano trzy różne modele sztucznych sieci neuronowych, które następnie poddano analizie porównawczej pod kątem efektywności ich działania. Każdy model został przygotowany w dwóch wariantach. Pierwszy wariant, dotyczący klasyfikacji, służył do przewidywania, czy wystąpią wstrząsy wtórne. Drugi wariant, dotyczący regresji, pozwalał określić czas wystąpienia oraz przewidywaną siłę tych wstrząsów.

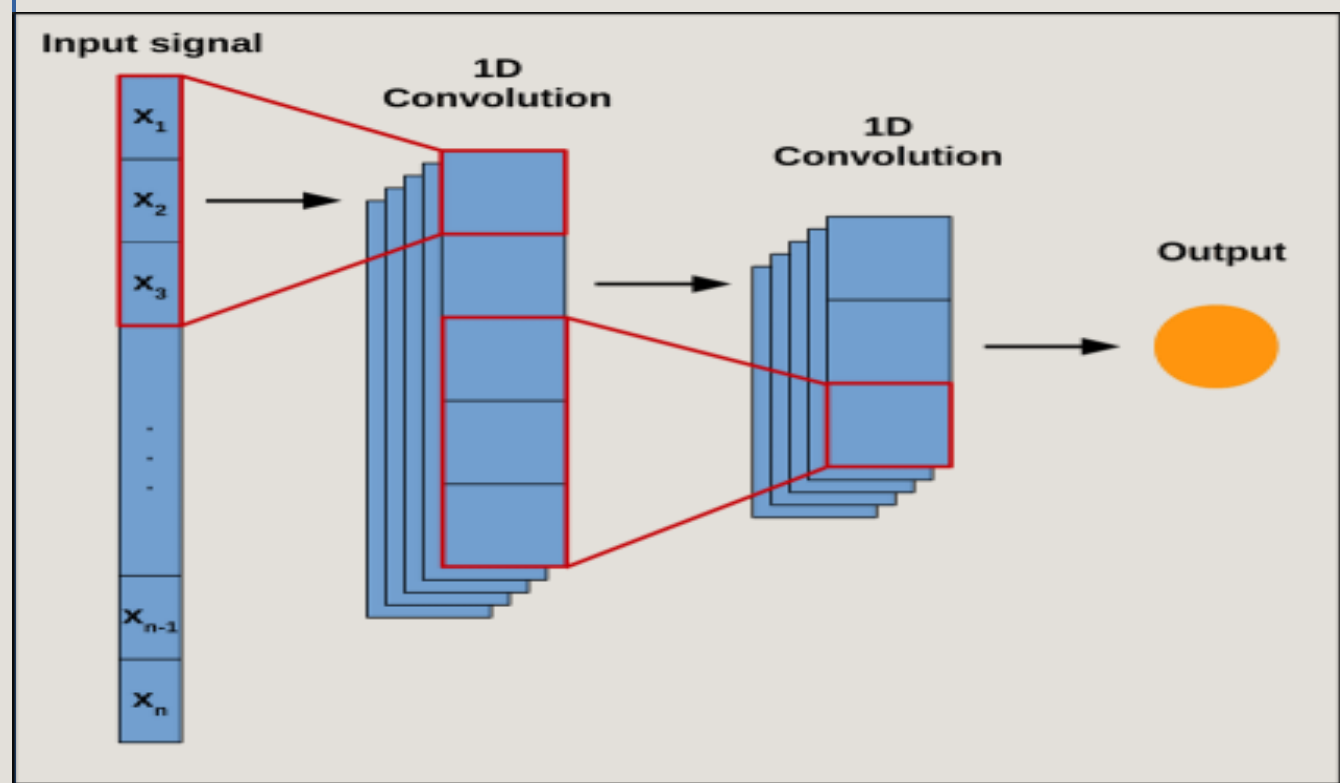
Wielowarstwowy perceptron

Wielowarstwowy perceptron (MLP) to klasyczna architektura sieci neuronowej złożona z co najmniej trzech warstw: wejściowej, jednej lub więcej ukrytych oraz wyjściowej. Każda para sąsiednich warstw jest w pełni połączona, a sygnały przepływają jednokierunkowo, bez sprzężeń zwrotnych. Dzięki nieliniowym funkcjom aktywacji (np. ReLU, sigmoida) MLP potrafi uczyć się złożonych, nieliniowych zależności między danymi wejściowymi a oczekiwanymi wynikami. Stosowany jest powszechnie w zadaniach klasyfikacji i regresji, gdzie pozwala na efektywną aproksymację dowolnych funkcji, o ile sieć ma odpowiednią liczbę warstw i neuronów.



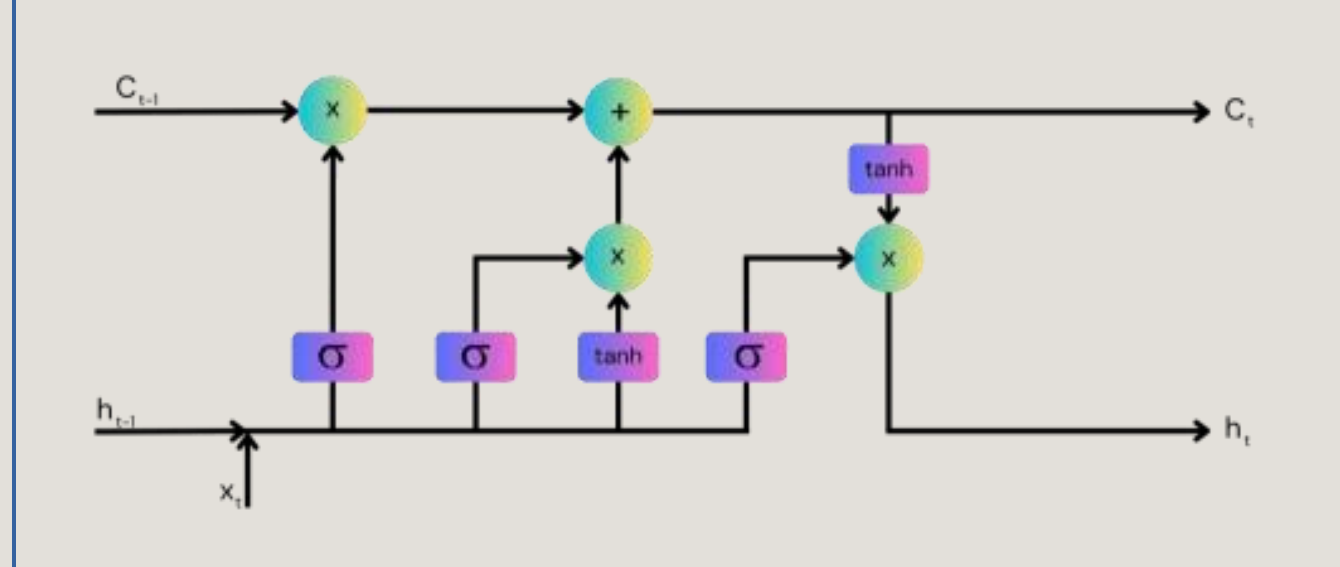
Sieci Konwolucyjne

Konwolucyjne sieci neuronowe (1D CNN) zostały stworzone przede wszystkim do realizacji zadań klasyfikacyjnych. W ramach niniejszego projektu zostały użyte w wersji jednowymiarowej (1D), ponieważ analizowane dane mają postać szeregu czasowego. Sieci te są szczególnie dobrze przystosowane do analizy szeregów czasowych, gdyż uczą się lokalnych wzorców czasowych za pomocą filtrów splotowych przesuwających się wzdłuż osi czasu. Podejście to efektywnie wychwytuje zależności krótkiego i średniego zasięgu w danych sekwencyjnych, a poprzez układanie kolejnych warstw możliwe jest rozszerzenie ich zdolności o naukę bardziej złożonych relacji. W konsekwencji sieci 1D CNN często zapewniają zarówno wysoką dokładność, jak i szybsze obliczenia w zadaniach takich jak prognozowanie.

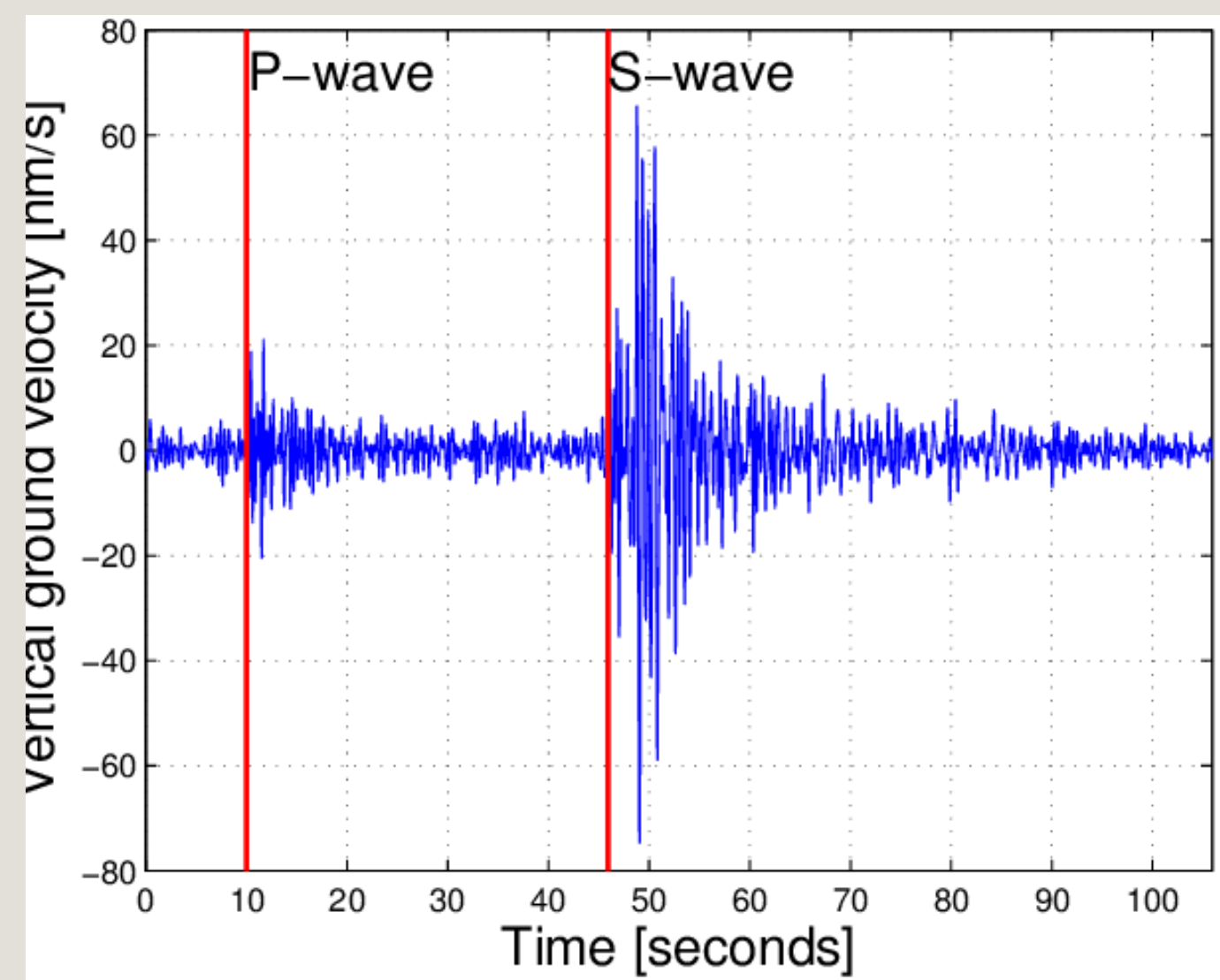


LSTM

Long Short-Term Memory (LSTM) to odmiana rekurencyjnych sieci neuronowych zaprojektowana do przechowywania długoterminowych zależności dzięki specjalnej architekturze komórek i trzech bramek (wejścia, zapomnienia, wyjścia), które kontrolują przepływ i modyfikację informacji. Z racji dużej ilości pomiarów w ramach jednego sejsmogramu, 6000 punktów w każdej z 3 orientacji, można go potraktować jako zapis długoterminowy.



Źródła:
[1] S. Mousavi, Y. Sheng "Stanford earthquake dataset (stead): A global data set of seismic signals for ai," IEEE Access
[2] P. Kavianpour, M. Kavianpour "A cnn-bilstm model with attention mechanism for earthquake prediction," The Journal of Super-computing,
[3]Q. Wang, Y. Zhang "On the use of vmd-lstm neural network for approximate earthquake prediction," Natural Hazards



Wyniki

W tabeli 1 ukazano wyniki dla sieci, które miały wskazać prawdopodobieństwo wystąpienia wstrząsów wtórnych. Z racji, że wyjściowy powinny wynosić 0 w przypadku nie wystąpienia wstrząsów wtórnych albo 1, gdy mają wystąpić, wartości 0,5 w metryce RMSE i MAE oznaczają, że sieć nie znalazła żadnych zależności w danych i zwraca losowe wartości. Sieć MLN osiągnęła zarówno Wskaźnik RMSE i MAE w okolicy 0.5, a R^2 jest ujemne. Jednak dla zbioru uczącego dane wartości wynoszą MAE: 0.1418 - RMSE: 0.2354 - R^2 : 0.7782. Różnice między wynikami na zbiorach uczącym i testowym jednoznacznie wskazują, że sieć jest przeuczona – model bardzo dobrze dopasował się do danych treningowych, ale traci zdolność do generalizacji na nowe przypadki, przez co na zbiorze testowym wyniki są losowe. W przypadku sieci CNN i LSTM osiągnęliśmy pewien stopień sukcesu, jednak metryki pokazują zbyt duży błąd, aby uznać je za skuteczne.

	RMSE	MAE	R ²
MLP	0,525	0,415	-0,102
CNN	0,354	0,251	0,498
LSTM	0.292	0.2025	0.622

Tabela 2 Pokazuje skuteczność w przewidywaniu czasu do wstrząsu wtórnego i proporcji jego siły do siły oryginalnego wstrząsu. W przypadku sieci MLN znowu pojawia się problem przeuczenia sieci. Jednak w wszystkich testowanych sieciach można zauważyć, że lepiej poradziły sobie z określeniem magnitudy wstrząsów wtórnych niż czasu do ich wystąpienia.

MLN			
	RMSE	MAE	R ²
Magnituda	0.617	0.529	-0.490
Czas	0.316	0.237	-0.606

CNN			
	RMSE	MAE	R ²
Magnituda	0,355	0,268	0,505
Czas	0,232	0,154	0,138

LSTM			
	RMSE	MAE	R ²
Magnituda	0,255	0,228	0,635
Czas	0,25	0,186	0,452

Pomimo użycia różnych sieci neuronowych, przewidywania wstrząsów na podstawie surowego sygnału sejsmicznego przyniosły gorsze wyniki niż sieci oparte na metadanych. CNN-BiLSTM [2] uzyskał R^2 równe 0,906 oraz RMSE i MAE 0,074. Natomiast VMD-LSTM osiągnął R^2 równy 0,9709, RMSE 0,0533 i MAE 0,0422.[3]