

Projekt ten rozpoczyna fascynującą podróż, mającą na celu odkrycie, jak różne czynniki naturalne i związane z działalnością człowieka wpłynęły na zanieczyszczenie powietrza, a konkretnie na stężenia cząstek PM_{2.5}, w całej Europie w ciągu ostatnich dwóch dekad. Cząstki PM_{2.5} to mikroskopijne zanieczyszczenia w powietrzu, które mogą mieć poważny wpływ na zdrowie, a zrozumienie ich źródeł i zachowania jest kluczowe dla poprawy jakości powietrza. Aby to osiągnąć, w naszych badaniach wykorzystamy nowoczesne techniki uczenia maszynowego do analizy dużych zbiorów danych z lat 2003-2023. Przyjrzymy się różnym czynnikom, w tym warunkom pogodowym, typom użytkowania gruntów, emisjom z różnych źródeł i innym zmiennym środowiskowym. Dzięki temu będziemy mogli przedstawić kompleksowy obraz, jak te elementy przyczyniały się do poziomów PM_{2.5} w czasie i w różnych miejscach na kontynencie. Główne założenie naszego badania opiera się na tym, że w związku ze zmianami klimatycznymi, a także ograniczeniem poziomów zanieczyszczeń w Europie w skali lat, wpływ niektórych naturalnych i antropogenicznych czynników na zanieczyszczenie powietrza mógł się zmienić.

Jednym z kluczowych narzędzi, które wykorzystamy, jest specjalna metoda zwana analizą wartości Shapleya (SHAP). Jest to nowatorskie podejście z dziedziny interpretowalnego uczenia maszynowego. Można go porównać do detektywa, który pomaga nam zrozumieć rolę każdego czynnika, wpływającego na przewidywanie stężeń PM_{2.5} w modelu. Ta metoda pozwala zobaczyć, które czynniki są najważniejsze w kształtowaniu poziomów zanieczyszczeń i jak ich wpływ zmienia się w różnych skalach czasowych. Na przykład, możemy odkryć, że niektóre czynniki pogodowe odgrywają kluczową rolę zimą, podczas gdy inne są bardziej istotne latem, lub że rola tych czynników zmieniała się w ciągu kilku lat. Nasze badania będą korzystać z wysokiej jakości danych z renomowanych źródeł, takich jak, m.in. Copernicus Atmosphere Monitoring Service (istniejące modele zanieczyszczeń), ERA5 (szczegółowa baza danych pogodowych), Corine Land Cover (baza o zagospodarowaniu terenu), WorldPop (dane o populacji), EMEP (dane o emisji zanieczyszczeń), oraz baza danych jakości powietrza Airbase. Integrując dane ze wszystkich źródeł, będziemy mogli stworzyć dokładne modele dokonujące przestrzennych przewidywań PM_{2.5}.

Aby upewnić się, że nasze wyniki są solidne, porównamy różne algorytmy uczenia maszynowego, w tym metody oparte na drzewach decyzyjnych i sieci neuronowe. Te algorytmy prezentują różne strategie lub podejścia do rozwiązywania problemu. Porównując je, możemy zidentyfikować, które z nich zapewniają najbardziej wiarygodne i spójne wyniki dla naszego badania. Choć ta porównawcza analiza nie jest głównym celem naszego projektu, zwiększy ona dokładność i wiarygodność naszych wyników. Sednem naszych badań jest zrozumienie trendów i wzorców wpływu różnych zmiennych na zanieczyszczenie PM_{2.5}. Chcemy wiedzieć, jak te wzorce różnią się nie tylko w trakcie doby oraz sezonowo, ale także na dłuższą metę, na przykład w ciągu całego dziesięciolecia. Dodatkowo, zbadamy, jak te trendy wyglądają w różnych częściach Europy, od tętniących życiem miast po ciche obszary wiejskie, zwracając też szczególną uwagę na warunki klimatyczne w tych regionach.

Pod koniec tego projektu spodziewamy się odkryć ukryte związki między zmianami klimatycznymi, redukcją emisji, dynamiką populacji a jakością powietrza. Wnioski wyciągnięte z tego badania będą wartościową informacją dla decydentów, naukowców i ogółu społeczeństwa. Upubliczniając nasze wyniki i zebrane dane, chcemy wspierać lepsze polityki i strategie w walce z zanieczyszczeniem powietrza oraz ochronie zdrowia publicznego w całej Europie. Podsumowując, projekt ten jest głęboką analizą skomplikowanej sieci czynników wpływających na zanieczyszczenie powietrza w Europie. Dzięki zaawansowanemu uczeniu maszynowemu i wysokiej jakości danym, zamierzamy ujawnić złożone wzorce i trendy dla zanieczyszczenia PM_{2.5} na przestrzeni ostatnich dwudziestu lat.