

Warsztaty uczenia maszynowego - projekt

Paweł Drabarek, Mateusz Fabisiewicz, Jakub Gazewski,
Adrian Goworek, Adam Wodzisławski

Marzec 2022

1. Temat i opis

Tematem projektu jest stworzenie rozwiązania wykorzystującego uczenie maszynowe służącego do analizy obrazów, na których mogą znajdować się ryby oceaniczne. Zadaniem programu jest zakwalifikowanie obrazu do jednej z 8 kategorii - 6 z nich dotyczy gatunku ryby, która znajduje się na obrazie (mogą to być – tuńczyk biały, tuńczyk wielkooki, koryfena, strojnik, gatunek rekina lub tuńczyk żółtopłetwy), pozostałe dwie odpowiadają sytuacji gdy na obrazie nie znajduje się ryba lub znajduje się ryba innego gatunku niż te rozpoznawane.

Szczegóły dotyczące danych znajdują się pod linkiem: www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring

2. Założenia technologiczne

Program zostanie wykonany w języku Python. Zostanie wykorzystana biblioteka umożliwiająca korzystanie z sieci neuronowej, prawdopodobnie PyTorch lub TensorFlow, nasze doświadczenie podczas pisania tego etapu dokumentacji nie pozwala nam jeszcze określić która z wymienionych bibliotek najlepiej pasuje do naszego problemu (być może ostatecznie wybrana zostanie zupełnie inna, nie wymieniona).

3. Podział pracy

W projekcie możemy rozróżnić takie etapy pracy jak:

- przygotowanie zbiorów treningowych i walidacyjnych;
- ustalenie odpowiedniego modelu, wyodrębnienie atrybutów i ustalenie hiperparametrów;
- implementacja i trening sieci neuronowej;
- testy modelu;
- sporządzenie dokumentacji.

Przy takim podziale etapów pracy, nasza grupa zostanie podzielona na następujące podzespoły:

- zespół I: przygotowujący aspekty teoretyczne naszego modelu;
- zespół II: implementujący rozwiązanie;
- zespół III: testujący i sporządzający dokumentację.

Osoby przypisane do zespołu będą zajmować się głównie daną częścią projektu, lecz przy braku zadań będą mogły zostać poproszone o pomoc w innych zespołach. Proponujemy również, aby niektóre osoby brały udział w pracach dwóch zespołów jednocześnie, umożliwi to szybsze i łatwiejsze zrozumienie owoców pracy innych zespołów.

Zespół I będzie miał za zadanie wykonać pierwsze 2 wypunktowane etapy pracy, zespół II zajmie się implementacją i treninigiem sieci, a zespół III testami uzyskanego modelu i spisaniem dokumentacji.

Do zespołu I należeć będą:

- Paweł Drabarek;
- Adrian Goworek;
- Jakub Gazewski.

W zespole II znajdują się:

- Jakub Gazewski;
- Adam Wodzisławski;
- Mateusz Fabisiewicz,
- Paweł Drabarek.

a w ostatnim zespole:

- Adam Wodzisławski;
- Adrian Goworek.