NORDIC MACHINE INTELLIGENCE

Nordic Machine Intelligence, vol. 02, pp. 1–3, 2022
https://doi.org/10.5617/nmi.9657

FishAI: Sustainable Commercial Fishing Challenge

Tor-Arne Schmidt Nordmo[1], Ove Kvalsvik[2], Svein Ove Kvalsund[2], Birte Hansen[3], Pål Halvorsen[4] , Steven A. Hicks[4] , Dag Johansen[1], Håvard Dagenborg Johansen[1], Michael A. Riegler[1,4]

1. *UiT The Arctic University of Norway, Norway*
2. *Vekstlandet, Norway*
3. *NORA—Norwegian Artificial Intelligence Research Consortium, Norway*
4. *SimulaMet, Norway*

Abstract

FishAI: Sustainable Commercial Fishing is the second challenge at the *Nordic AI Meet* following the successful MedAI, which had a focus on medical image segmentation and transparency in machine learning (ML)-based systems. FishAI focuses on a new domain, namely, commercial fishing and how to make it more sustainable with the help of machine learning. A range of public available datasets is used to tackle three specific tasks. The first one is to predict fishing coordinates to optimize catching of specific fish, the second one is to create a report that can be used by experienced fishermen, and the third task is to make a sustainable fishing plan that provides a route for a week. The second and third task require to some extend explainable and interpretable models that can provide explanations. A development dataset is provided and all methods will be tested on a concealed test dataset and assessed by an expert jury.

Keywords: artificial intelligence; machine learning; fishing; automatic reporting

Introduction

With a fishing zone spanning 2.1 million square meters, Norway is considered Europe's largest fishing and aquaculture nation. Every year, commercial vessels catch fish with a total value of around 20 billion NOK from the Norwegian fishing zone.

The overall migration patterns of the major fish species are relatively predictable and common knowledge. A fisherman knows, for example, that the mackerel season starts in mid-September and plans accordingly. On a daily basis, however, fish populations can move over large distances, and with the main decision-making tool being the captain´s experience and intuition, boats often search for days or even weeks before making a catch. The number of boats is not negligible; there are currently around 1,100 Norwegian vessels over 11 meters involved. It is estimated that each vessel burns around 2,000–2,500 liters of fuel per day, which translates to approximately 5 million kg CO2-equivalents per day.

Although the fishing fleet over time has shown an impressive ability to renew itself, the core operation of searching and catching fish clearly has room for improvements in a sustainability context. Specifically, a more energy-efficient commercial fishing practice and operation should be a goal. In other words, there are great environmental benefits and opportunities in optimizing commercial fishing activities by reducing unnecessary transport distances. With the recent release of catch data made available by the Norwegian Directorate of Fisheries, a significant potential of applying artificial intelligence opened up, which we want to explore with this challenge.

Dataset Details

We provide the participants with a collection of four publicly available datasets: a catch note dataset, a temperature dataset, a salinity dataset, and a moon phase dataset. All datasets can be used in all tasks and can be downloaded via: https://tinyurl.com/54w5bvxa. For the GPS coordinates predictions, the catch notes dataset is the ground truth. Participants are also encouraged to use other data sources if they are public available. In the following we provide a more detailed description of each dataset and what the participants can expect for the evaluation of their results.

Catch Notes Dataset

The catch notes data contains catch notes collected by the Norwegian Fishing Directorate from 2000 to today for vessels larger than 15 meters. The notes consist

© 2022 Author(s). This is an open access article licensed under the Creative Commons Attribution License 4.0. (http://creativecommons.org/licenses/by/4.0/).

of information about the catch that is manually logged during landing, e.g., when it was caught, where it was caught, what equipment was used, and the species distribution of the catch. There are approximately 130 data fields and around one million notes each year. Fields that might be of interest include information about where they fished ("Hovedområde", "Lon", "Lat", etc.) and information about fish caught and how they were caught ("Art - FDIR", "Bruttovekt", "Redskap").

The catch notes are in Norwegian. For most of the variables this is not relevant. In case it might be relevant participants can translate the data as part of the their data preparation pipeline. The dataset from each year can be found, along with documentation and metadata, at: https://www.fiskeridir.no/Tall-og-analyse/AApne-data/Fangstdata-seddel-koblet-med-fartoeydata.

Temperature, Salinity, and Moon Phase Datasets
The temperature, salinity, and moon phase datasets are meant to be auxiliary datasets that might give more information regarding fish migration. Both the temperature and salinity datasets are presented in netCDF4 formats. Therefore, we recommend to use the `netCDF4` Python module for extracting the data.

Temperature Data

Sea surface temperature (SST) from 1981 to present has been collected by National Oceanic and Atmospheric Administration (US). It contains daily estimates of SST globally. The data was collected from satellite observations, and consists of daily data at 0.25 degree latitude $\times$ 0.25 degree longitude resolution [1]. We have included the subset of data from 2000 to present day. The dataset is published at: https://www.psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html.

Salinity Data

Monthly averages of salinity data from 2015 to present day is provided from the SMAP Salinity V4 dataset [2]. Salinity (in combination with temperature) affects the growth rate of microalgae. This can potentially affect the migration patterns of fish. Eight-day running averages are also possible to obtain if needed (https://salinity.oceansciences.org/data-smap-v4.htm).

Moon Phase Data

The moon phase data consists of dates and exact times of full moon from 1900 to 2050. Lunar phases affect the migration and behaviour of fish due to water levels changing. Therefore, it is potentially possible to use this data source for modeling of the movement of fish. The dataset is published at https://www.kaggle.com/datasets/lsind18/full-moon-calendar-1900-2050.

Task Descriptions

We present three subtasks: the catching optimization task, the reporting task, and the planning task. Each task targets different aspects of the data. The participants are encouraged to submit solutions for all three subtasks, but can also just focus on specific tasks.

Task 1: Catching Optimization and Prediction Task
Build a model that can predict which coordinates a vessel should prioritize in order to maximize the likelihood of catching a type of fish of your choosing (haddock or mackerel is most valuable for the industry partners). The prediction can be based on historical data.

Task 2: Report Generation Task
Create a report of your analysis that can be read by experienced fishermen; an user-friendly visualization that a captain can read to make a assessment of where the vessel should search for fish the next day.

Task 3: Sustainability Fishing Plan Task
Make a Sustainability Fishing Plan; a weekly plan that suggests the routes the fisherman should follow to optimize fish caught and fuel consumption.

The aim is to build a tool that will help fisherman make decisions about where to search for fish in the immediate future.

This could include features such as a heatmap indicating the largest likelihood for catch of a specific type of fish, recommendations based on predicted catch volume relative to distance from current location, or similar.

Evaluation Methodology

Task one will be evaluated using an unseen test set. We will use standard metrics such as precision, F1-score, accuracy, mean absolute error, etc., to evaluate the performance of the methods. Tasks two and three will be evaluated by an expert team consisting of experienced fishermen and data scientists which will provide an overall ranking of the submitted report and fishing plan. Submissions to task two and three are evaluated using a qualitative approach compared to task one. The quality of the report and plan are measured by attributes like readability, presentation, and usefulness. Each team will receive a report from the expert team on their performance. There will be one first place and one second place based on a combination of the evaluations gathered from each of the three sub-tasks.

Summary

This article presents the *FishAI: Sustainable Commercial Fishing* challenge held at the 2022 Nordic AI Meet. The challenge aims to open up for research and innovation in the commercial fishing domain to increase sustainability. FishAI is providing three subtasks that range from catch optimization to automatic report generation. We hope that this challenge inspires established and young researchers and people interested in innovation to explore

an important and interesting topic contributing to more sustainable commercial fishing activities.

Conflict of interest

There is no conflict of interest.

References

1. Reynolds RW, Smith TM, Liu C, Chelton DB, Casey KS, and Schlax MG. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. Journal of Climate 2007; 20:5473–96. DOI: 10.1175/2007JCLI1824.1. Available from: https://journals.ametsoc.org/view/journals/clim/20/22/2007jcli1824.1.xml
2. Meissner T, Wentz FJ, and Le Vine DM. The Salinity Retrieval Algorithms for the NASA Aquarius Version 5 and SMAP Version 3 Releases. Remote Sensing 2018; 10. DOI: 10.3390/rs10071121. Available from: https://www.mdpi.com/2072-4292/10/7/1121