Ataskaitinė informatikos inžinerijos krypties doktorantų konferencija

Ataskaita už 2018 – 2019 mokslo metus

Doktorantas: Julius Venskus

Vilnius, 2019

Bendra informacija

- **Disertacijos pavadinimas:** „Giliojo mašininio mokymo taikymas jūrų transporto eismo duomenims klasifikuoti“
- **Darbo vadovas:** dr. Povilas Treigys
- **Konsultantas:** dr. Arūnas Andziulis
- **Doktorantūros pradžia:** 2016 m.
- **Planuojama doktorantūros pabaiga:** 2020 m.

Informacija apie tyrimą

- **Tyrimo objektas:**
  - jūrų transporto priemonių eismo duomenys
- **Tyrimo tikslas:**
  - ištirti ir pasiūlyti būdus jūrų transporto duomenų klasifikavimui ir neįprasto eismo aptikimui.
- **Tyrimo uždaviniai:**
  - apžvelgti esamus jūrų transporto klasifikavimo ir neįprasto eismo aptikimo būdus AIS (automatinės identifikavimo sistemos) duomenyse;
  - ištirti klasikinius jūrų transporto klasifikavimo ir neįprasto eismo aptikimo AIS duomenyse būdus;
  - ištirti ir pritaikyti giliuoju mokymusi grindžiamus algoritmus transporto duomenims klasifikuoti;
  - atlikti eksperimentinį klasifikatorių vertinimą.
- **Planuojami rezultatai:**
  - jūrų transporto eismo klasifikavimo ir neįprasto eismo aptikimo algoritmus įgyvendinantis programų sistemos prototipas.

2018/2019 darbo planas

STUDIJŲ PLANAS

- **Mokslinių tyrimų publikavimas:**
  - Empirinio tyrimo rezultatų publikavimas recenzuojamuose leidinyje.

- **Dalyvavimas konferencijose, seminaruose, kitose doktorantų mobilumo veiklose:**
  - Disertacijos empirinio tyrimo rezultatų pristatymas tarptautinėje konferencijoje.
  - Disertacijos išvadų pristatymas tarptautinėje konferencijoje

Ataskaita už 2018/2019 mokslo metus (1)

- **Mokslinių tyrimų publikavimas:**
- **Indeksuojami WOS periodiniai leidiniai**
  - Venskus, Julius; Treigys, Povilas; Bernatavičienė, Jolita; Tamulevičius, Gintautas; Medvedev, Viktor. **Real-time maritime traffic anomaly detection based on sensors and history data embedding //** Sensors. Basel : MDPI. ISSN 1424-8220. 2019, vol. 19, no. 17, art. no. 3782, p. 1-10. DOI: 10.3390/s19173782. Q1
- **Recenzuojami leidiniai**
  - Venskus, Julius; Treigys. **Meteorological Data Influence on Missing Vessel Type Detection Using Deep Multi-Stacked LSTM Neural Network** // XII International Conference Computer Data Analysis & Modeling 2019 Stochastics & Data Science. Proceedings of the XII International Conference. September, 18-22, Belarusian State University, 2019 Minsk, Belarus, ISBN-978-985-566-811-5
  - Venskus, Julius; Treigys. **Preparation of training data by filling in missing vessel type data using deep multi-stacked LSTM neural network for abnormal marine transport evaluation**. Proceedings of Abstracts. ITISE 2019 International Conference on Time Series and Forecasting. Granada, Spain, September, 25-27, 2019, ISBN 978-84-17970-79-6

Ataskaita už 2018/2019 mokslo metus (1)

- **Pranešimai konferencijose**
- **Stendiniai:**
  - Venskus, Julius; Treigys. Preparation of training data by filling in missing vessel type data using deep multi-stacked LSTM neural network for abnormal marine transport evaluation. ITISE 2019 International Conference on Time Series and Forecasting. Granada, Spain, September, 25-27, 2019

- **Žodiniai:**
  - Venskus, Julius; Treigys. Meteorological Data Influence on Missing Vessel Type Detection Using Deep Multi-Stacked LSTM Neural Network // XII International Conference Computer Data Analysis & Modeling 2019 Stochastics & Data Science. September, 18-22, Belarusian State University, 2019 Minsk, Belarus

Ataskaita už 2018/2019 mokslo metus (2)

- **Dalyvavimas seminaruose, kitose doktorantų mobilumo veiklose:**
  - Venskus, Julius; Treigys, Povilas; Bernatavičienė, Jolita; Andziulis, Arūnas. **Aspects of data collection for abnormal marine transport evaluation** // DAMSS 2018 : 10th international workshop on "Data analysis methods for software systems", Druskininkai, Lithuania, November 29 - December 1, 2018 : [abstract book]. Vilnius : Vilniaus Universitetas, 2018. ISBN 9786090700433. p. 88.
  - Pranešimas: ***VU Duomenų mokslo ir skaitmeninių technologijų instituto KSG seminare***: Venskus Julius, Savi-organizuojančio neuroninio tinklo (SOM) ir virtualaus feromono integravimas neįprastam laivų eismo aptikimui, 2018m. Vilnius

Ataskaitinė informatikos inžinerijos krypties doktorantų konferencija

Mokslinė ataskaita

Doktorantas: Julius Venskus
Vadovas: dr. Povilas Treigys
Konsultantas: prof. dr. Arūnas Andziulis

Vilnius, 2019

VILNIUS UNIVERSITY
INSTITUTE OF DATA SCIENCE AND DIGITAL TECHNOLOGIES

Meteorological data influence on missing vessel type detection using deep multi-stacked LSTM neural network

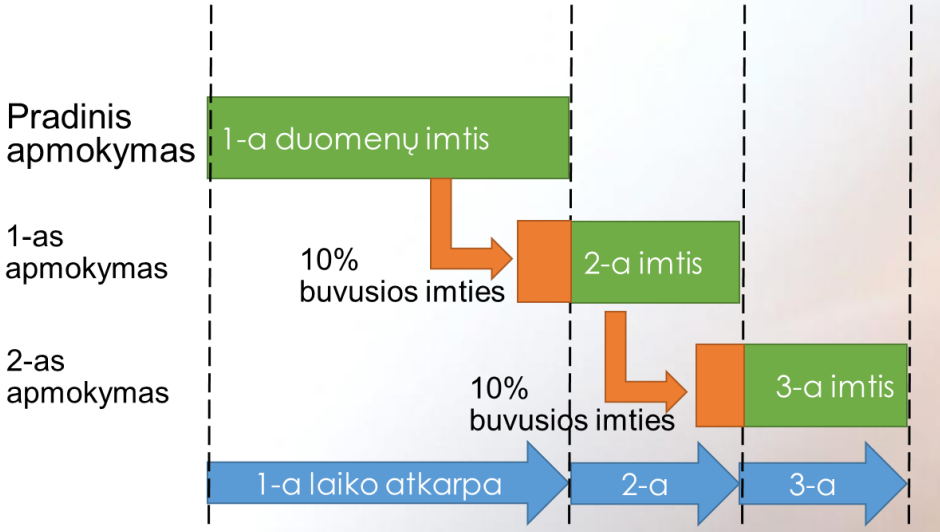
PhD candidate: Julius Venskus
Supervisor: Dr. Povilas Treigys

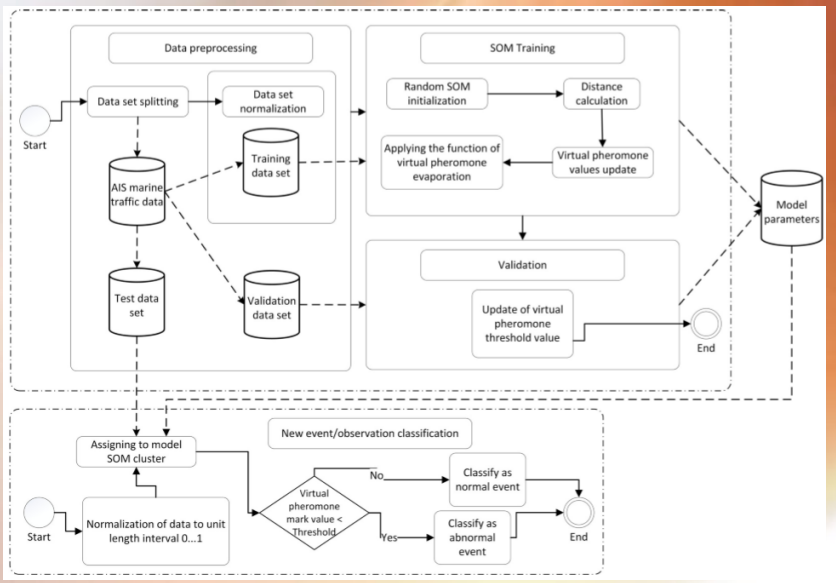
Vilnius, 2019

Continuation of 2016/2017/2018 work

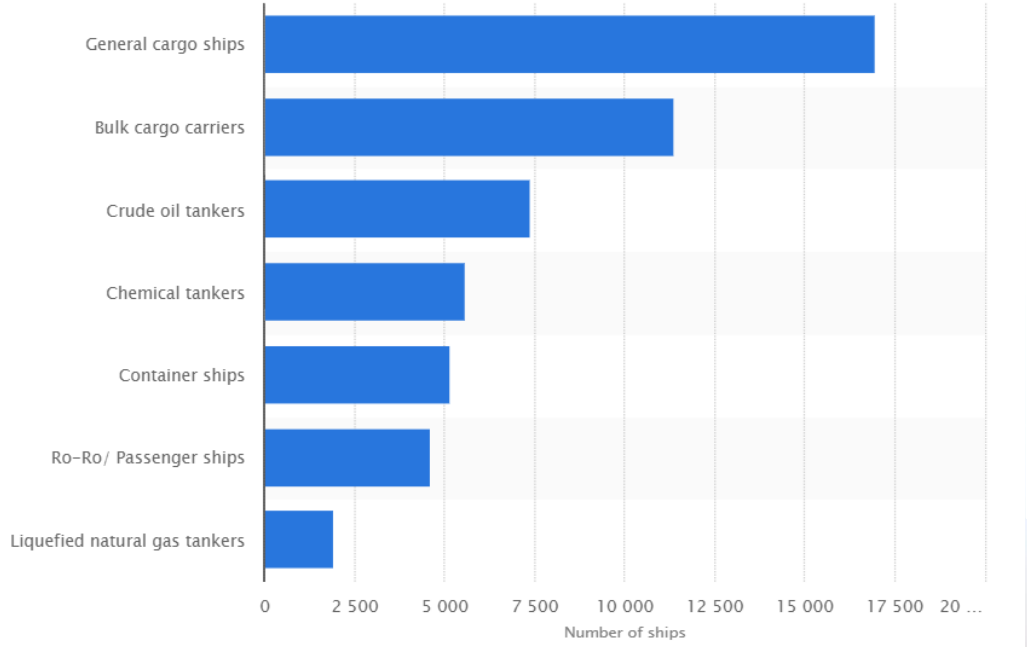
2016/2017 created algorithm for vessel abnormal movement detection based on SOM with integrated virtual pheromone and compared to M.Riveiro SOM-GMM algorithm.

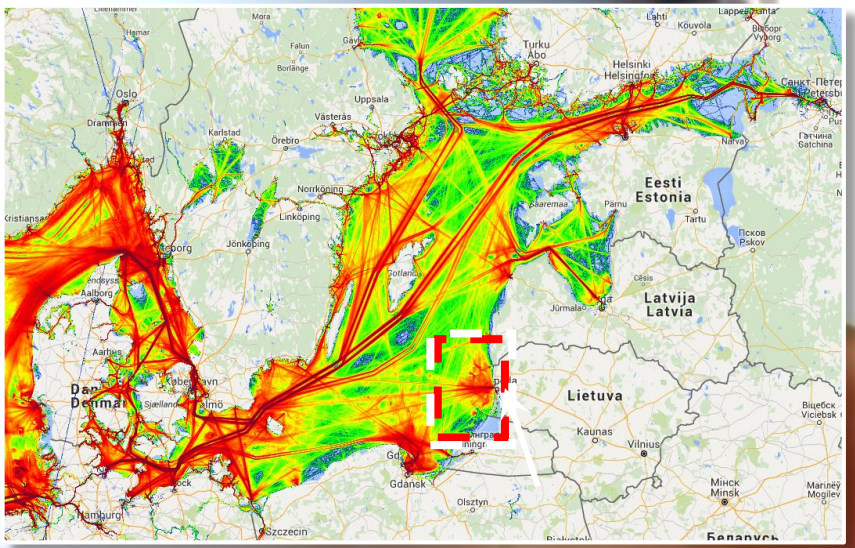
Unsupervised Neural Network Retraining: Real-time Maritime Traffic Anomaly Detection Based on History Data Embedding

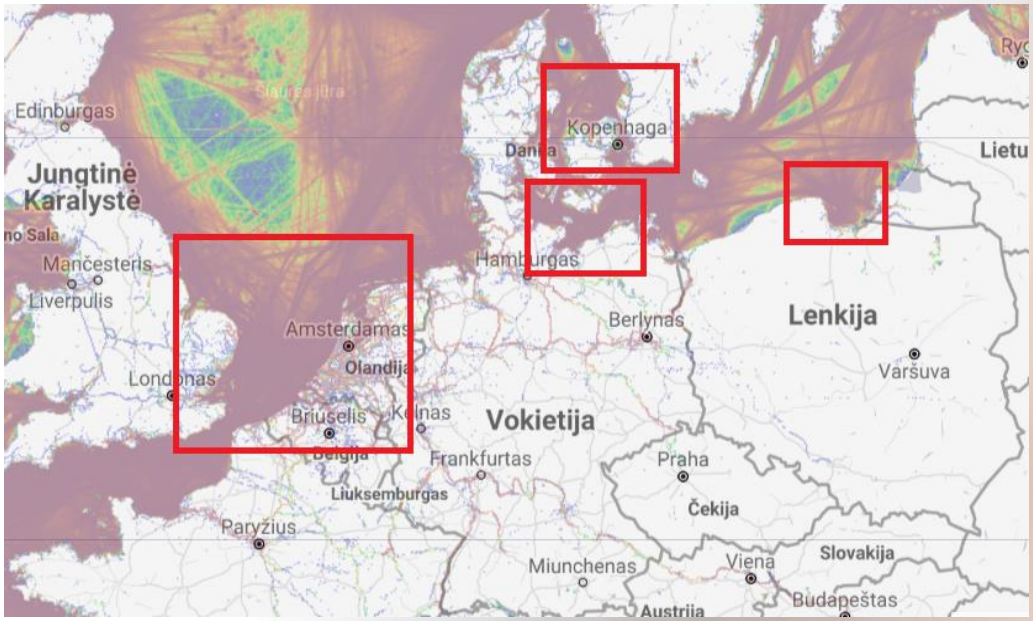




Big data of marine vessel traffic







- Marine transport – 90% of total trade
- High risk of marine traffic incident
- Earthly detection of abnormal activity gives opportunity minimise accident's risk
- The world's merchant fleet as of January 1, 2018, with a breakdown by type. Of the around 53,000 merchant ships trading internationally

Maritime traffic anomaly

Khatkhate et al.(2007) use the following definition within mechanical systems: *[a]n anomaly is defined as deviation from the nominal behavior of a dynamical system and is of-ten associated with parametric and non-parametric changes that may gradually evolve in time*."

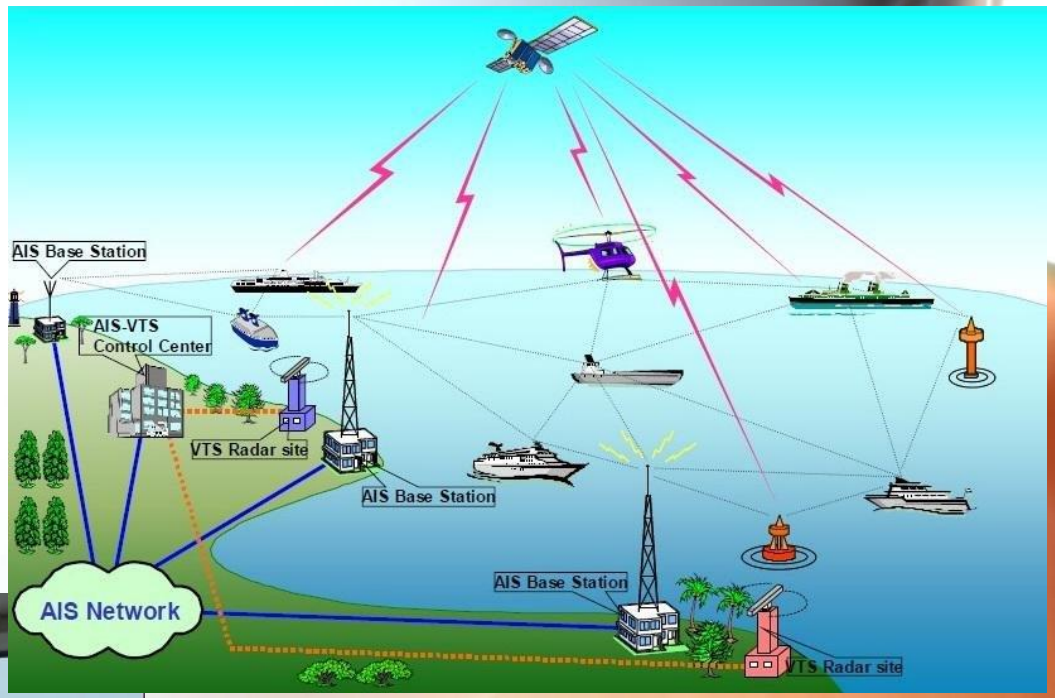
Goals of detection:

- Vessels collision.
- Vessel traffic into shallow water or other obstructions
- Vessel malfunction
- Vessel hijack
- Smuggling
- Espionage or reconnaissance
- Piracy
- Illegal fishing
- Military manoeuvres
- Territorial violations

AIS,VTS and missing data

VTS – Vessel traffic service. Control and navigational services. Rescue initiation.

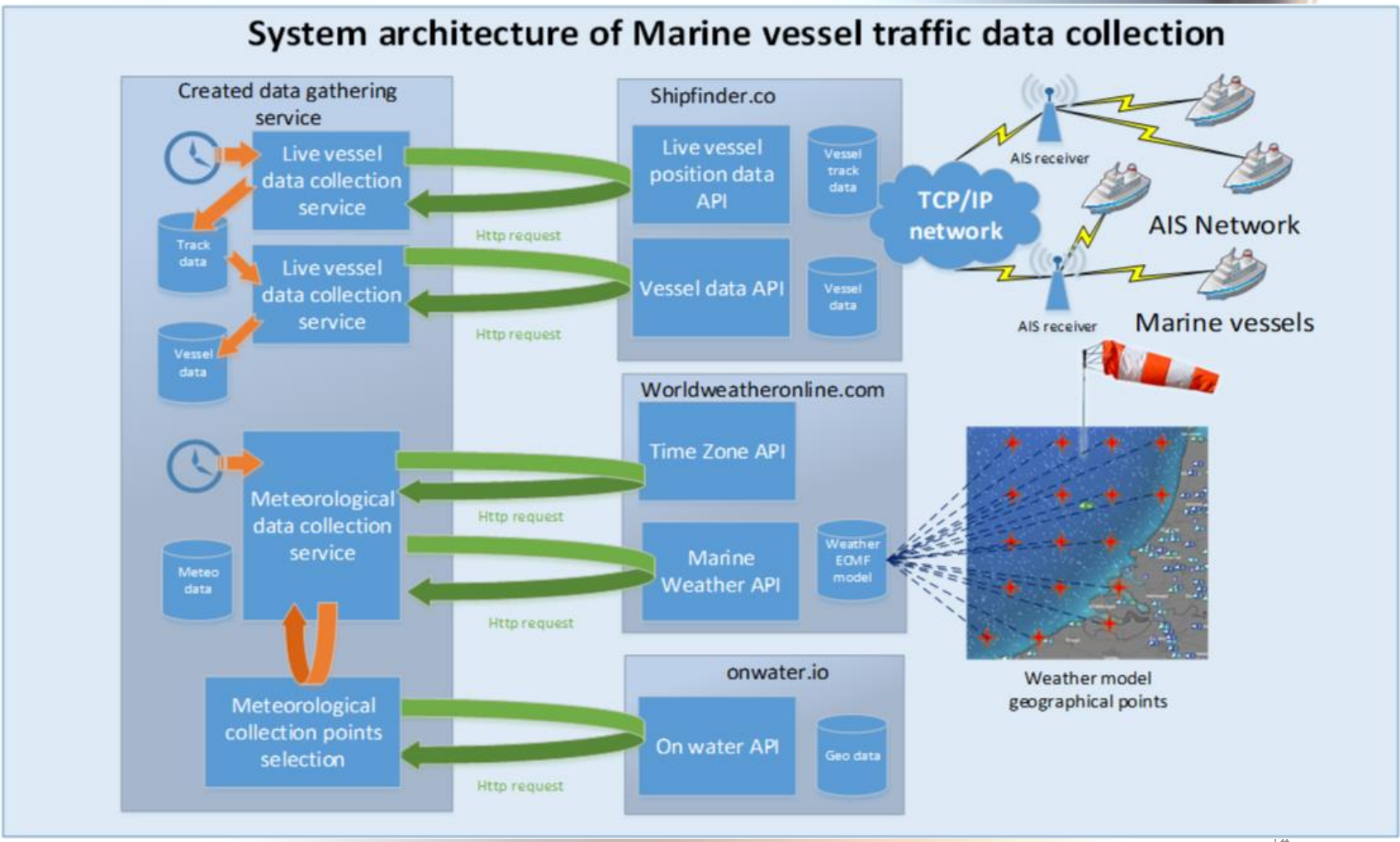
AIS – Automatic identification system. Vessel ID, position, speed, heading, status, port call, vessel type and etc.



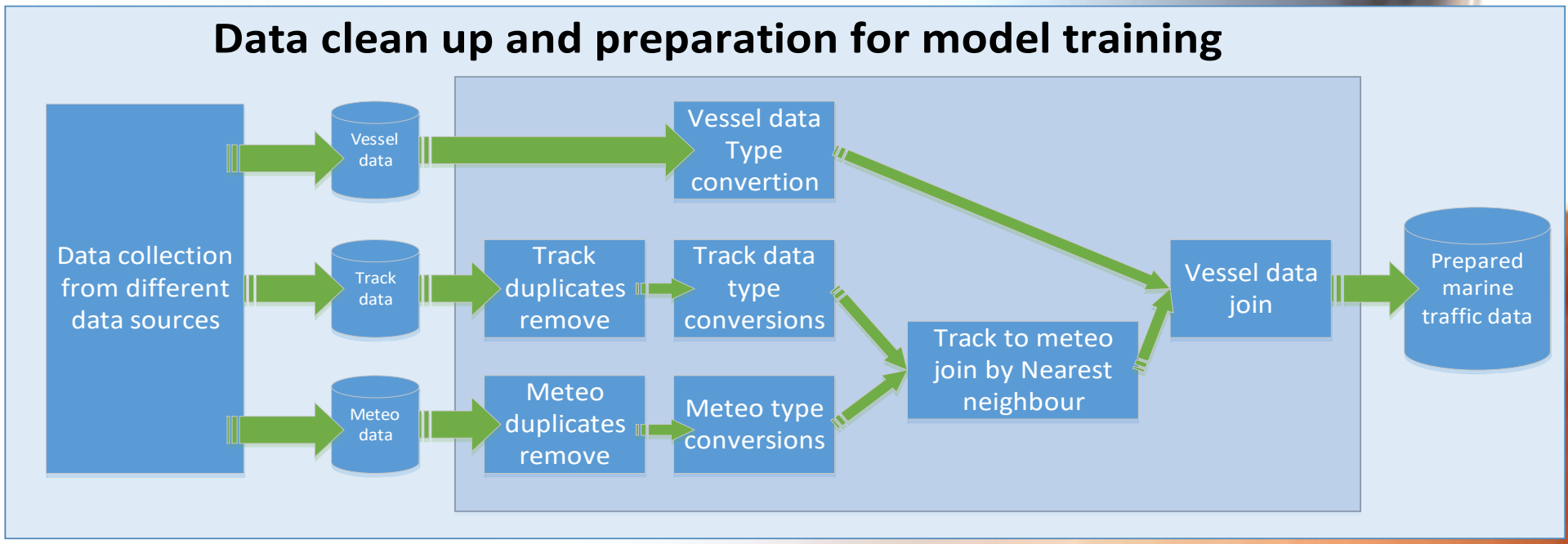
Missing a vessel type information in marine traffic data

Depending of data source, missing

Data collection from multiple open data sources



Data preparation



Assigning a meteorological data location to vessel track point by Nearest neighbour

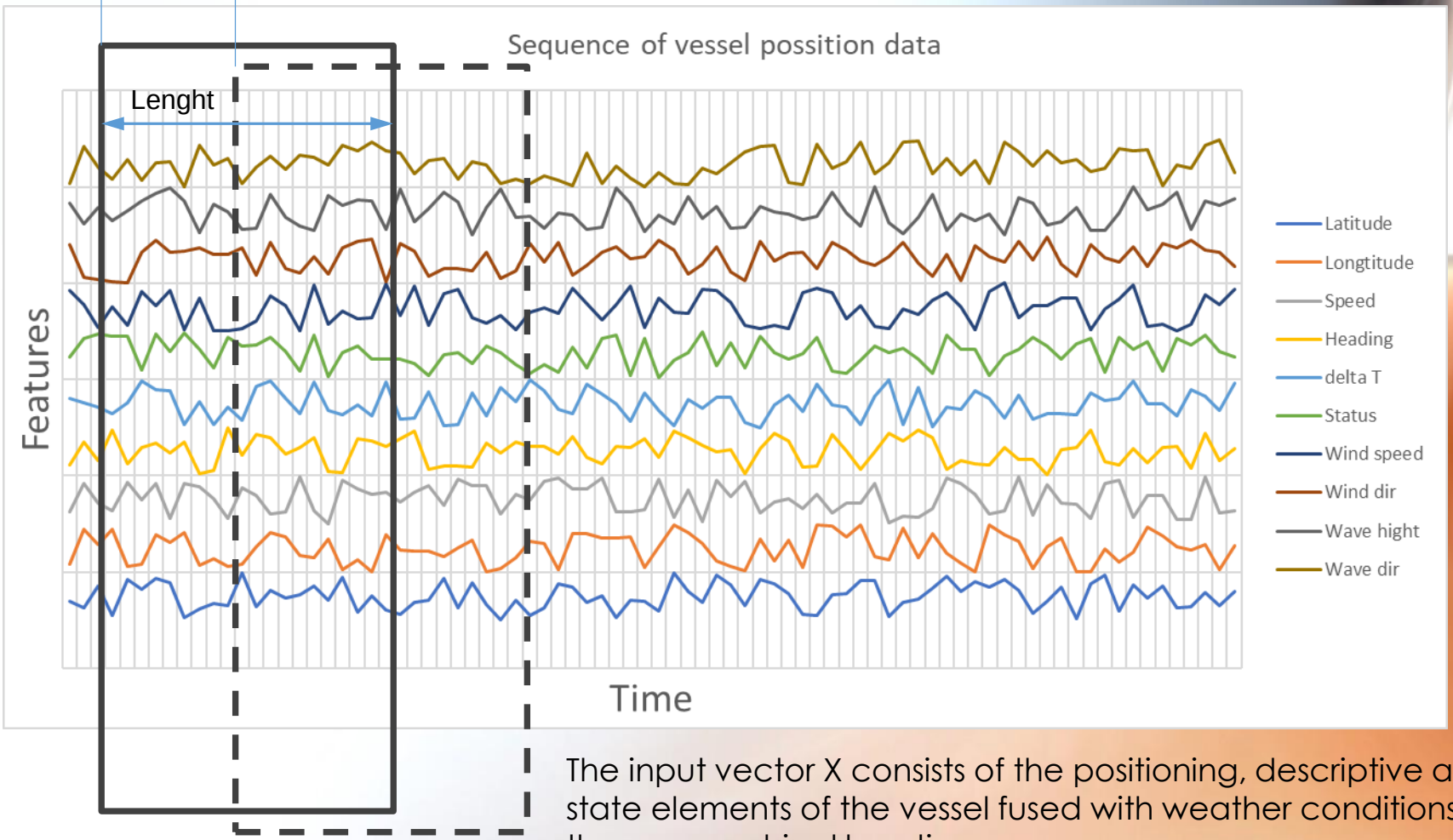
Vessel track point

Vessel path line

Meteorological data location

Association of Meteorological data location to vessel track piont

Forming data sequences



Sliding sequence frame

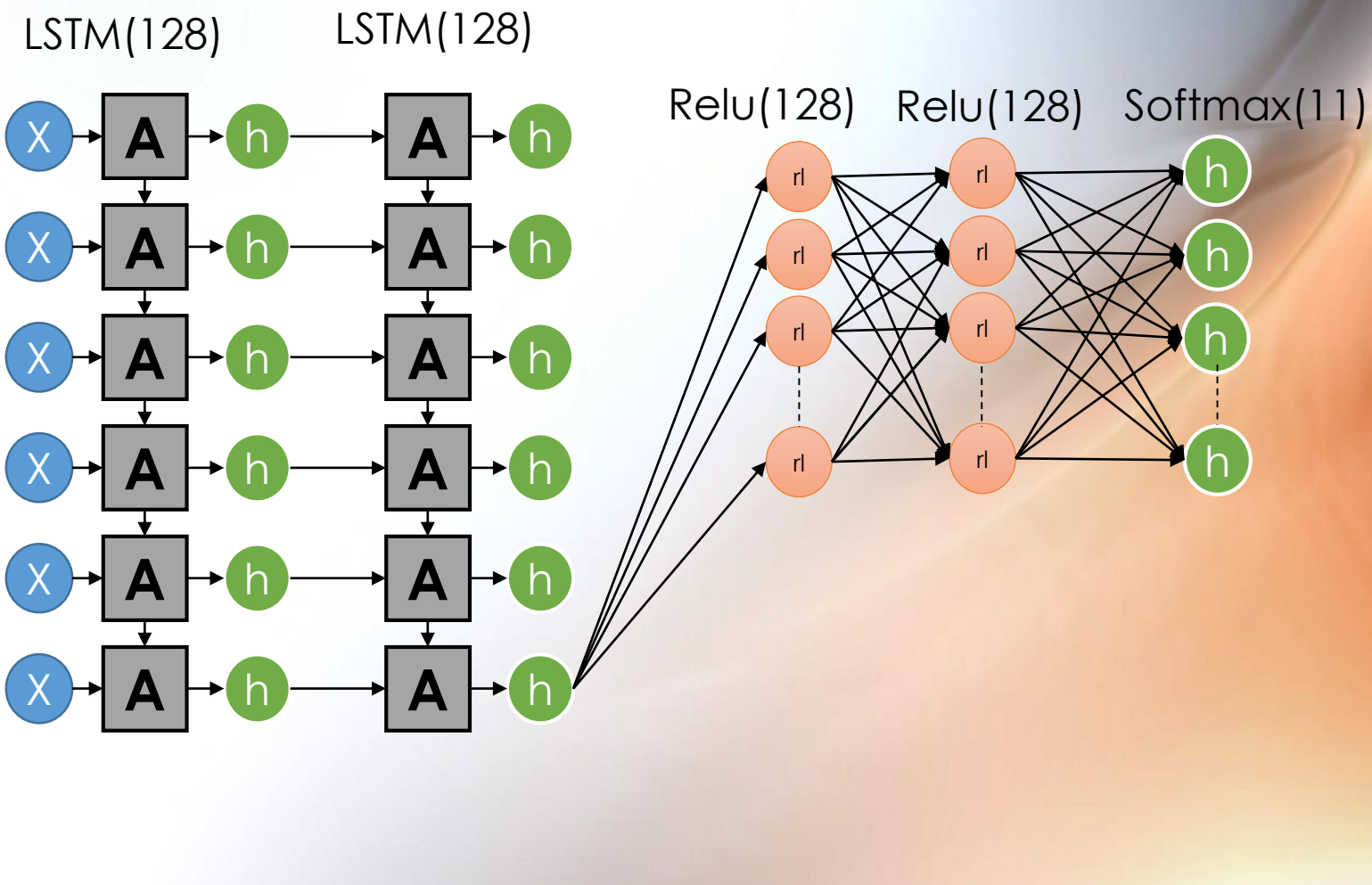
The input vector X consists of the positioning, descriptive and state elements of the vessel fused with weather conditions in the geographical location.

$$X_T^p = \begin{bmatrix} X^1 \\ X^2 \\ \vdots \\ X^p \end{bmatrix} = \begin{bmatrix} X^1_{T-(n-1)} & X^1_{T-(n-2)} & \cdots & X^1_{T-1} & X^1_T \\ X^2_{T-(n-1)} & X^2_{T-(n-2)} & & X^2_{T-1} & X^2_T \\ \vdots & & \ddots & \vdots & \\ X^p_{T-(n-1)} & X^P_{T-(n-2)} & \cdots & X^p_{T-1} & X^p_T \end{bmatrix},$$

Main parameters:

- Step
- Stride
- Sequence length

Architecture of stacked LSTM deep neural network



Validation on real data set

Geographical region - Rotterdam harbor area, The Netherlands
Data time period – November 1, 2018 to November 30, 2018

Vessel traffic data collected:

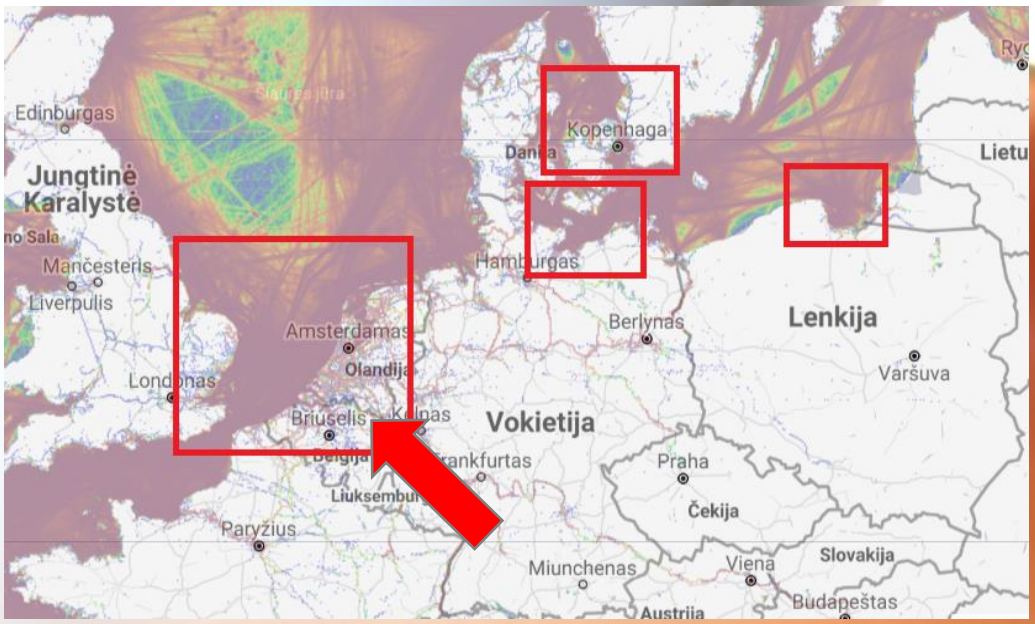
	Data vectors	Persentage
Vessel type available	1.20E+06	4.12%
Missing vessel type	2.78E+07	95.88%
Total	2.90E+07	100%

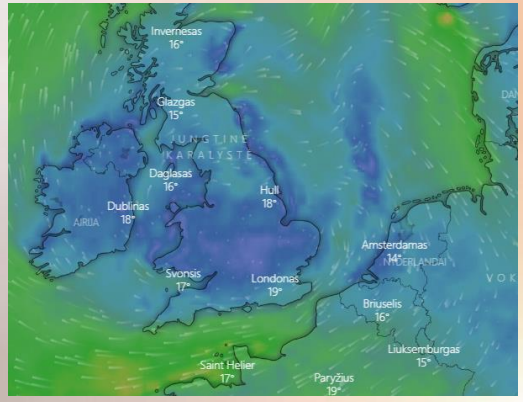
Meteorological data collected:

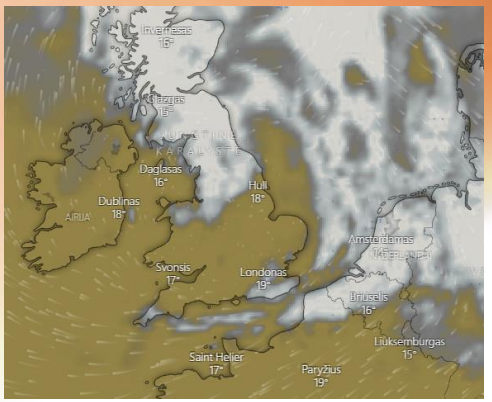
Area	Astronomical data vectors	Marine weather data vectors
Vector quantity	3000	18000
Vector size	0.4kB	3kB

Data split for training and testing:

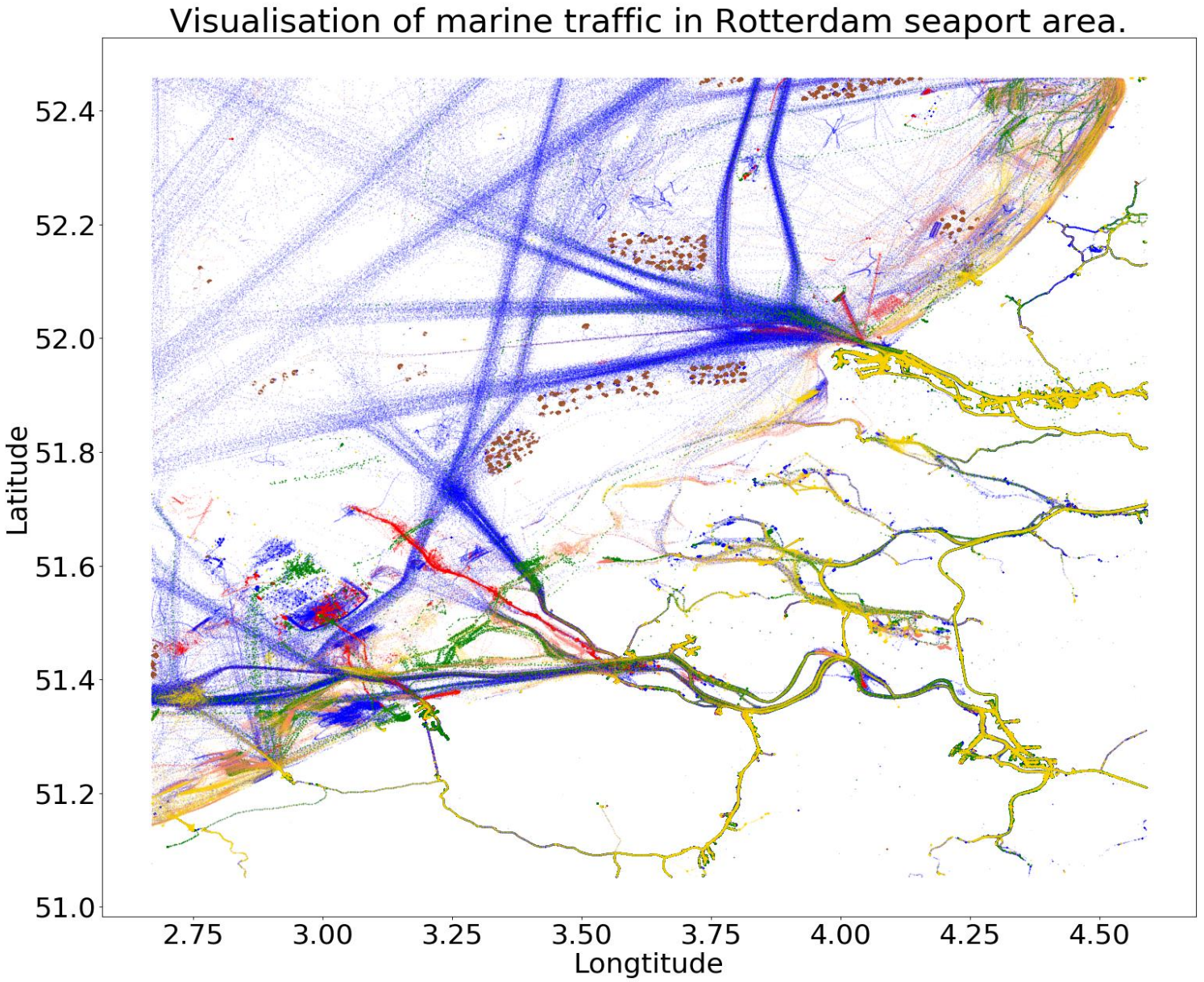
	Sequences of vessel position data	Persentage
Training set	600000	50%
Validation set	355000	30%
Test set	240000	20%
Total	1195000	100%





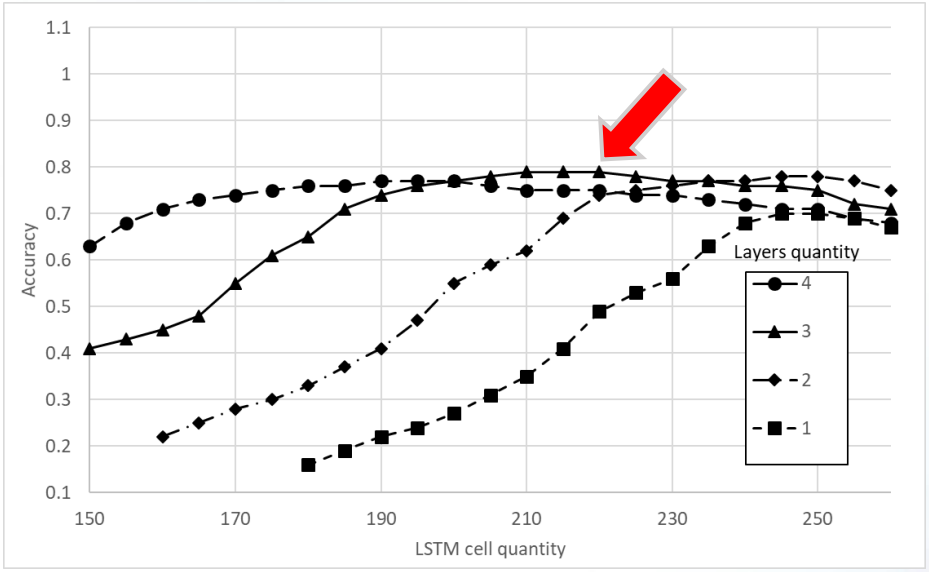


Visualization of marine traffic



Results

Only vessel traffic data:



Meteorological data included:

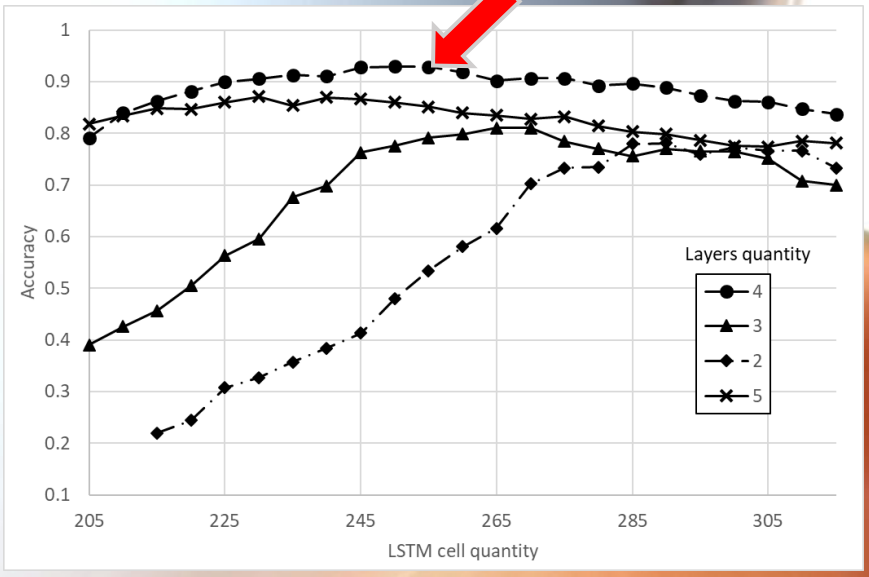


Table 1: Trend of classification accuracy for different network settings

	Meteorological data excluded		Meteorological data included	
Layers	Cells	Accuracy	Cells	Accuracy
2	245	0.78	290	0.78
3	215	0.79	265	0.81
4	195	0.77	250	0.93

Future work

Marine vessel movement prediction based on available position data.

Anomalous marine vessel traffic detection based on deep LSTM network.

Thank you

Ačiū
Klausimai?