“The Use of Big Data Analytics for Process Modelling in Smart Logistics Operations”

*Call Topic*: Big data and process modelling for smart industry (BDSI), Call 2017

*Start date*: 01/04/2019

*End date*: 31/03/2022
BIG-SMART-LOG Partners

- Ekol Lojistik
- Centre for Research and Technology Hellas (CERTH) / Information Technologies Institute (ITI)

- Lublin University of Technology
- Holisun SRL
- Antalya Bilim University
The project aims to address the challenge of redesigning the entire data produced by:
- satellite enabled (IoT) vehicle tracking technology (traffic, weather and road condition information, sensor data with vehicle/load information)
- customers’ orders
- third party logistics services (carriers, logistics and terminal operators, etc.),
- open-source systems
The goal of the project is to design a semantic-enhanced self-learning processing model which:

- analyzes operational data in real-time
- uses learned data to influence the utilization of existing infrastructure and resources
- increases system resiliency and service quality
- reduces greenhouse gas emissions, fuel consumption, idle times, traffic congestion, revenue losses, late deliveries, customer complaints

Using technologies such machine learning, and deep learning algorithms that will provide predictions to:

- maximize operational efficiency,
- environmental protection and safety

The outputs of the project will produce a reusable framework for dynamic environments with customizable constraints and parameterized data sources.
### Main Objectives

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop planning systems for freight to find the best (combinations of) modes and optimal route to progress towards freight traceability information allowing better network exploitation and more efficient logistics operations.</td>
<td>10% shorter delivery routes</td>
</tr>
<tr>
<td>Having the routes and checkpoint locations optimized, the organization can achieve positive environmental impact of the service via reduced average fuel consumption per parcel.</td>
<td>Decrease 15%</td>
</tr>
<tr>
<td>Providing real time information such as traffic jams, weather and road conditions.</td>
<td>Very high</td>
</tr>
<tr>
<td>Reduction in driver turnover, driver assignment.</td>
<td>Decrease 15%</td>
</tr>
<tr>
<td>Reducing the time spent in transit and idle time.</td>
<td>Decrease 20%</td>
</tr>
<tr>
<td>Operational efficiency will be capitalized into increased quality of service to the final customers; from the sender’s part, average response time is expected to decrease.</td>
<td>Decrease 30-40%</td>
</tr>
</tbody>
</table>
Project’s Structure / Work Packages Dependencies
Summary of Achievements

- Collection of business use cases, data and requirements
- Architecture design
- Logistics ontology design
- Literature reviews
- Data pre-processing and initial analysis results
- Scientific publications
- Participation in international events
Business Use Cases

- **UC1**: Long term forecasting of time and location of future orders

- **UC2**: Long term forecasting of time and location of pickups and deliveries

- **UC3**: Forecasting and decreasing late deliveries
• **Scientific literature review** on data mining/machine learning algorithms/methods on:
  • relevant works approaching issues concerning logistics operations
  • publications related to the identified business use cases

• **Categorization** based on machine learning tasks:
  • classification approaches
  • regression approaches
  • hybrid methods

• **Algorithms and methods:**
  • *Time delay or arrival prediction:*
    • Classification and Regression Trees (CART),
    • Logistic Regression,
    • Support Vector Regression (SVR),
    • Random Forest etc.
  • *Location prediction:*
    • Content-based methods (like Markov model),
    • Distribution-based methods,
    • Pattern-based methods etc.
Literature review of algorithm’s evaluation criteria

• **Scientific literature review** on evaluation criteria of the machine learning algorithms

• **Categorization** based on the most common practices of evaluation of machine learning models/tasks:
  - classification
  - regression
  - clustering

• **Criteria and metrics:**
  - Pearson’s, Mathews Correlation coefficient
  - Precision, Recall, Accuracy and F-measure
  - Receiver Operating Characteristic (ROC) Curve and its Area Under the ROC Curve (AUC)
  - Precision – Recall Curve (P-RC) and its Area Under Curve (AUPR)
  - (Normalized) Root Mean Square Error (RMSE, NRMSE)
Available Data

- Ekol Lojistik provided data concerning their freight transshipment operations.

- Describing transportation of orders to/from different customers and destinations in Europe, through trucks, trains and vessels.

- Data exploratory analysis for extracting correlations and useful information.
A logistics ontology has been created in order to semantically describe the processes and data related to project use cases.

The ontology contains over 100 classes and sub-classes containing major means related to logistics such as:

- Logistic operations and services
- Vehicles and containers
- Transportation modes
- Location and distances
- Orders
- Operators
Framework Architecture

Data Management: Retrieving and storing data based on following sub-components

- **Live Data Aggregator**: Sensors, GPS etc.
- **Database Query Engine**: Historical data stored in database and querying mechanisms
- **Open Data Collector**: Data from public sources

Data Pre-processing: Turn data into a united, compatible form, ready for our main processing

Data Analysis Framework: The "brain" of the whole system

- **Training Module**: Feed existing data to our models and improve their accuracy
- **Analytics Module**: Exploratory data analysis to reveal hidden insights

Decision Support: Suggest decisions and actions, based on the analytics outcome

- **Optimization Engine**: Help optimize the process, reduce costs, make best choices
- **Visualization Module**: Visual representations of the data analytics, graphs, plots etc.
High Level BD Visual and Data Analytics tool Architecture

Functional View Diagram
Dissemination

- Continuous dissemination through project’s webpage and social media accounts
- Participation in 9 international events
- 14 scientific publications to peer reviewed journals and proceedings of international conferences
Next Steps

- Implementation and evaluation of methods and algorithms
  - Determining appropriate algorithms according to problem and objective definition
  - Implementing of the machine learning components
  - Evaluating techniques, methods and algorithms

- Development of integrated data analytics web platform
Thank you for your attention