METHOD TO LABEL STOP LOCATIONS’ PURPOSE

The Analytics and Architecture group at Scania Connected services and Solutions drives the advancement of applied data science and machine learning within connected services. A thesis project in our group is a great opportunity to work with applied machine learning at the forefront of the transport industry and an excellent way of making contacts for your future career. Help us map the transport system – stop locations, their purpose and how they connect!

Background
Scania is one of the world’s leading manufacturer of trucks and buses for heavy transports, as well as industrial and marine engines. Transport services and logistics services make up an increasing part of our business, which guarantees Scania’s customers cost-efficient transport solutions and high availability. Over a million Scania vehicles are in active use, in over 100 countries.

At Connected Services & Solutions we develop new solutions for connected vehicles in our Internet of Things platform, as part of Scania’s increasing focus on communication, services and smart transport solutions. Advanced data analysis capabilities are a cornerstone enabler in this development.

Description
Scania has more than 300 thousand connected vehicles continuously sending in position information. This data is core to insights of the trucks and the transport system.

A label, describing the purpose of the stop location, could for example be loading, parking, unloading or refueling. Given a set of purpose labels we want to map out the locations and label them. With this map we can discover logistic patterns, optimize transport flows and gain a deeper understanding of how the trucks are used. Increased understanding of stop locations enables the future development of services to better conform with Scania’s customers’ needs in their daily operations. Ultimately increased understanding will support the development of a sustainable transport system and smart cities.

Your thesis work will be carried out in a team of data scientists and architects with diverse competencies giving you first-hand insight into how we at Scania drive our business from data. You will not only have access to data sets with billions of observations and big data computation platforms, but you will also have access to our Apache Spark Streaming code pipeline and ground truth data enabling you to evaluate your methods. This thesis’ already has potential customers and our ambition is that you will cooperate with them during the development.

Assignment
Develop a method to label stop locations. This will include building an understanding of manually created purpose labels and create a more generic set of these labels, designing and extracting features from a set of stops, explore and evaluate various approaches and finally establish a method with potential customers.

Applicants
Education: Master’s program in machine learning, data science, computer science or similar. Applicants are expected to have a good understanding of relevant machine learning and data mining methods.

Meritorious: Experience of Spark Streaming, programming in Scala/Java and addressing big data problems. We love working with data, data processing tools, big data platforms/frameworks and challenging problems. We always aim to create value for the business as well as the society at large, and we want applicants to consider themselves driven by the same intentions.

Number of students: 1-2
Start date: Flexible
Estimated time need: 6 months

Applicants will be assessed on a continuous basis until the position is filled.

See more on scania.com

Apply here!