# Jori van Lier – Freelance Data Scientist


| | |
|---|---|
| Location: | Leiden, The Netherlands |
| Gender: | Male |
| Nationality: | Dutch |
| Email: | info@jvlanalytics.nl |


I create business value using data science and machine learning. I enjoy building data products with some sort of predictive element to help people make important decisions. I'm highly proficient with Python, Hadoop and Spark. As a team player with 7 years of professional experience, I function well in a senior role in a team, where I can take up a bit of project management and mentor junior data scientists.

## Work experience

### Oct 2017 – Onwards: Freelance Data Scientist

Creating business value using data science and machine learning.

### Oct 2016 – Sep 2017: KPMG Advisory – Senior Data Scientist

I was responsible for analytics at BlueSense, a start-up that was incubating within KPMG. BlueSense is a platform that monitors the behaviour of people within buildings, in real-time, using a lot of sensors. It includes state of the art analytics to understand crowd density, crowd movement patterns, dwell times and occupancy levels, both descriptive and predictive. This system was deployed at various clients in retail, public transport, facility management and a football stadium.

### Dec 2013 – Sep 2016: KPMG Advisory – Senior Consultant Big Data & Analytics

I was part of a team of big data experts with a mission to disrupt industries with novel data-driven approaches. As one of the pioneers in the big data space, we brought machine learning, simulations and statistics to regular businesses. I did a mix of client projects and product development for a real-time Location Analytics platform (which would later evolve into BlueSense).

### 2010 – 2013: KPMG Advisory – (Senior) Consultant IT Advisory

Business intelligence and data quality projects. Senior Consultant as of 2012.

### 2007 – 2010: Webtechniek – Developer (part time)

Developing various PHP/MySQL web applications.

# Education

### 2008 – 2010: Leiden University – MSc Computer Science

Focus on artificial intelligence algorithms (neural networks, evolutionary algorithms, genetic programming, swarm intelligence). Graduated cum laude.

### 2004 – 2008: Leiden University – BSc Computer Science

### Various courses

Deep Learning Specialization, deeplearning.ai on Coursera, 2017 (in progress).

Management Development, KPMG, 2017.
Calculus Two, Ohio State University on Coursera, 2017.
3-day advanced statistics course, Nikhef (the Dutch institute for subatomic physics), 2016.
Calculus One, Ohio State University on Coursera, 2016.
Machine Learning, Stanford University on Coursera, 2014.
Statistics, Duke University on Coursera, 2014.

# Skills and Competences

**Languages:** Dutch (native), English (fluent).

**Programming languages:** Highly proficient with Python and Java. Average knowledge of JavaScript. Familiar with the basics of Scala. Frequent user of the following Python packages: NumPy, pandas, matplotlib, scikit-learn, SciPy, Flask, PyMC3, Keras, TensorFlow.

**Big Data platforms**: Hadoop (Hortonworks), Hive, YARN, HDFS, Spark, Storm, Kafka.

**Web:** NodeJS, JavaScript, Python Flask, Rest APIs, HTML, CSS, Bootstrap, JQuery, a little D3.js.

**Databases:** SQL in various dialects, mostly MySQL, PostgreSQL and Hive. NoSQL: MongoDB.

**Dev tools:** Git, Maven, Archiva, Jenkins, Docker, Vagrant, JIRA.

**Linux:** CentOS, Ubuntu, Bash scripting, Amazon Web Services.

**Machine learning:** Most of my experience involves linear regression, logistic regression, (stochastic) gradient descent, decision trees, random forests, boosted trees, k-means clustering, (convolutional) neural nets, backprop, PCA, k-fold cross-validation, learning curves, ROC curves, bias-variance dilemma, ensembles.

**Statistics**: Familiar with both frequentist and Bayesian approaches. I take a simulation based approach to frequentist hypothesis testing and uncertainty estimation. Good understanding of probability theory, likelihood, confidence intervals, $chi^2$ tests, EM, Bayes rule, Bayesian credible intervals, Bayes factor. Moderate understanding of Bayesian methods involving MCMC.