

# Vasilis Vryniotis

Data Scientist & Software Engineer

🏠 datumbox.com | 📄 datumbox | 🌐 bbriniotis | 🐦 bbriniotis

---

## Summary

I'm a Data Scientist, a Software Engineer and the main author of Datumbox Machine Learning Framework. I have worked for companies such as Expedia, King.com, Microsoft & Mozaik and I have co-founded WebSEAnalytics.com. My studies include an MSc in Machine Learning, an MSc in Statistics, an MBA and a BSc in Computer Science.

I am particularly interested in Machine Learning; both in its mathematical foundations and its applications. I specialize in building complex Machine Learning Frameworks & Pipelines, developing custom Statistical Models & Algorithms and performing Data Modeling. Moreover I have experience in working in the areas of Computer Vision, Natural Language Processing, Recommender Systems, Deep Learning, Learn-to-Rank and Real-Time Bidding.

---

## Work Experience

### Expedia Group - Principal Data Scientist

11.2017 - Present

As a Principal Data Scientist, I head up the team responsible for Computer Vision, Natural Language Processing and Recommender Systems of Hotels.com. Additionally I lead the R&D efforts of Expedia in the area of Deep Learning and I participate in the design and development of the Machine Learning systems required for training and deploying models on the live environment.

⇒ Key Technologies: Machine Learning, Deep Learning, Computer Vision, NLP, Recommender Systems, Python, Scala, Java, Keras, TensorFlow, Spark, SQL, Hive, Hadoop.

### Expedia Group - Head of Data Science

11.2015 - 10.2017

Expedia, Inc. is one of the largest online travel companies in the world, with an extensive brand portfolio that includes some of the world's leading online travel brands (Hotels.com, Trivago etc). The main responsibilities of my role included leading the Data Science team, conducting R&D for the development of novel algorithms, building Machine Learning models for the Search Engine and Recommender Systems of the company, developing Machine Learning pipelines and working on various data modeling projects.

⇒ Key Technologies: Machine Learning, Learn-to-Rank, Recommender Systems, NLP, Computer Vision, Deep Learning, Scala, Java, Python, Spark, Keras, TensorFlow, SQL, Hive, Hadoop.

### King - Senior Data Scientist

09.2014 - 11.2015

King.com is the leading interactive entertainment company for the mobile world; some of the most well-known games of the company include the Candy Crush, Farm Heroes and Pet Rescue Saga. The key responsibilities of my role included developing Statistical & Machine Learning models for the Real-Time Bidding System (DSP/advertising) of the company, building

Recommender Systems, developing Machine Learning pipelines in Python & Java, designing and running complex A/B-tests and writing Business Reports.

⇒ Key Technologies: Machine Learning, NLP, Python (Numpy, Scipy, Scikit-Learn, Pandas, Graphlab), Java, SQL, Hive, Hadoop, Mahout, QlikView.

### **Web SEO Analytics - Co-Founder & Technical Director**

08.2009 - 08.2013

WebSEOAnalytics.com, a corporate spin-off of Mozaik, was among of the most well-known SEO & Online Marketing tool providers with major clients from all over the world. As the co-Founder & Technical Director of the company, I was leading the team responsible for the Analytics Platform and I had hands-on involvement in its design and development. Moreover I was leading the efforts to improve the infrastructure of our backend systems in order to store and access efficiently our multi-TB Search Engine Index. Finally I was responsible for the development of the Machine Learning models which powered up our Analytics Platform and were used for Ranking Prediction, Text Extraction, Spam Detection, Sentiment Analysis and other Text Classification and Natural Language Processing applications.

⇒ Key Technologies: Machine Learning, NLP, Web Crawling, Java, MySQL, MongoDB, Apache Pig, Web Services.

### **Mozaik - Head of Software Development**

09.2007 - 07.2013

Mozaik.com is a leading multi-awarded web development & online marketing agency which specializes in the Tourism & Online Retail industries and has offices in Athens, Paris and Nicosia. As head of the development department I was managing directly a team of 9 software engineers and a system administrator. I was involved in the design & development of internal tools, responsible for analyzing and managing large web projects and heavily involved in the coding of Web Services, CMS Systems, Portals, E-commerce Platforms, Payment Systems, Online Booking Systems and Online Marketing tools.

⇒ Key Technologies: PHP, CakePHP, MySQL, JSON, XML, HTML5, CSS3, Web Services.

### **Microsoft - Software Engineer**

03.2007 - 06.2007

Microsoft Hellas and the Institute of Technology in Greece proposed the adoption of the Shibboleth Single Sign-On System in all the Greek Universities in order to build a secure network between organizations and protect the exchanged personal information. The system was initially designed for Microsoft's MSDNAA service and was later extended to support other 3rd party providers. In this contract project I was responsible for the analysis and development of the portal & web service and for the creation of the secure network in the 3 largest universities of Greece (University of Piraeus, National Technical University of Athens, Kapodistriako University of Athens).

⇒ Key Technologies: C#.NET, ASP.NET, Microsoft SQL Server, Shibboleth Server.

### **Freelancer - Web Developer & Software Engineer**

05.2006 - 03.2007

As a freelance Web Developer and Software Engineer I worked for numerous Web Development and Software Development companies. My main responsibilities were the analysis, design and development of Online Services, Portals, Websites and Windows Desktop Applications.

⇒ Key Technologies: PHP, C#.NET, ASP.NET, Microsoft SQL Server, MySQL, PostgreSQL.

---

## Projects

### **Datumbbox Machine Learning Framework**

07.2013 - Present

I am the author of Datumbbox Machine Learning & Statistical Framework. Datumbbox is written in Java, open-sourced under Apache License and capable of handling large data. It supports the most popular Machine Learning algorithms used in Classification, Regression, Cluster Analysis, Dimensionality Reduction & Feature Selection and the most important Statistical methods used in Statistical Tests, Data Exploration, Sampling & A/B testing. To get the source code or details on the supported methods visit: <https://github.com/datumbbox/datumbbox-framework>

Datumbbox.com is a free Web Service which uses the Framework to provide Machine Learning & Natural Language Processing algorithms. It offers a powerful REST API which can be used in a broad spectrum of applications including: Sentiment Analysis, Topic Classification, Language Detection, Subjectivity Analysis, Spam Detection, Reading Assessment, Keyword & Text Extraction and more. To get more information on the supported APIs visit:

<http://www.datumbbox.com/>

### **Open-Source & Blogging**

08.2009 - Present

I am actively involved in the development of many open-source projects and over the years I have contributed code to several well-known frameworks and applications such as Keras, Spark, Dstat etc. You can find more details on my Github account: <https://github.com/datumbbox>

I am the main author of Datumbbox.com Blog where I publish articles on Machine Learning, Statistics, Natural Language Processing, Text Analysis and Big Data. Moreover I have been a regular author on WebSEOAnalytics.com Blog where I published articles primarily about Analytics & SEO.

---

## Education

### **Imperial College London - MSc in Machine Learning**

09.2013 - 09.2014

### **Athens University of Economics and Business - MSc in Statistics**

09.2011 - 09.2013

### **Athens University of Economics and Business - MBA**

09.2009 - 09.2011

### **University of Piraeus - BSc in Computer Science**

09.2005 - 09.2009

---

## Technical Skills

<b>Data Science</b>	Machine Learning, Statistics, Computer Vision, Natural Language Processing, Recommender Systems, Deep Learning, Learn-to-Rank, Real-Time Bidding.
<b>Data Tools</b>	Spark, Keras, TensorFlow, Python Notebooks, IBM SPSS Statistics, Mahout, Datumbox, Graphlab, Microsoft Excel.
<b>Languages</b>	Scala, Java, Python, SQL.
<b>Big Data Systems</b>	Spark, Hive, Hadoop, Apache Pig.
<b>Database Systems</b>	MySQL, PostgreSQL, Microsoft SQL Server, MongoDB.
<b>Technologies</b>	AWS, Web Services, Unit Testing, Git, Bash.

---

## Languages

<b>English</b>	Full professional proficiency.
<b>Greek</b>	Native proficiency.
<b>French</b>	Elementary proficiency.