

# Santosh Tirunagari *March 7, 1988*

tsantosh7@gmail.com • +44 (0) 7914961317 • tsantosh77 (Skype) • <https://www.linkedin.com/in/santosh-tirunagari/>

---

## Summary

I'm an experienced Data Scientist with a PhD in AI / machine learning, with 8+ years background in predictive analytics, data driven modelling, data visualisation, multivariate data analysis, feature extraction, natural language processing (NLP), computer vision (CV), software / web development and cloud computing.

My present work at Synoptica focuses on developing NLP based novel software-as-a-service AI tools, which enable sales and marketing teams to save their time and sell their products better and faster. The AI tools I developed are available at <https://goo.gl/4uLb87>.

Previously, I was a Research Fellow in "machine learning for healthcare" at the University of Surrey (MRC) - developing ML based AI algorithms and software aimed at improving understanding of chronic diseases such as diabetes and Chronic Kidney Disease (CKD). These models, validated by expert clinicians, give doctors access to actionable predictions about their patients' chronic conditions. For this work, I was recognised with the **CogX 2018 award**: <https://goo.gl/QHnDpA>.

My PhD was about machine learning, signal processing and computer vision, and was fully sponsored from the University of Surrey (87,000 GBP). During my PhD, I was recognised with three awards including the "Excellent Oral Presentation" from IEEE ICIVC <https://goo.gl/pQaZU5>, <https://goo.gl/PBJm8w> and "Best Research Potential" and "Outstanding Service" awards <https://goo.gl/nbZWoR>.

I have significant experience in Python, JQuery, MATLAB, Apache Solr/Lucene, SQL, Linux, GIT, Agile, AWS, L<sup>A</sup>T<sub>E</sub>X, & web/software application development. I also write grant proposals and research articles <https://goo.gl/5gdZ4U>.

In general, I have a broad experience of machine learning, AI and data science across a variety of domains as well as significant leadership and mentoring experience in these areas. Furthermore, I have commercial software engineering experience of putting prototypes into production, liaising with development and front end teams and just making it work.

---

## Experience

### Synoptic Technologies Ltd.

COUNTY DURHAM, UNITED KINGDOM

#### Data Scientist

*Jul '17 – Present*

Work involves developing artificial intelligence algorithms using state-of-the-art machine learning techniques that help B2B companies automate research and prioritise leads based on company indicators made up of deep web and proprietary data sources. This enables sales and marketing teams to save time and sell faster. The AI tools I developed are available at <https://goo.gl/4uLb87>. Moreover, I contributed to a proposal, that attracted GBP 50,000 grant from Innovate UK to develop an AI based web application using Machine Learning for effectively managing their data.

*Programming & Tools:* Python, TensorFlow, Keras

*Methods :* NLP, SVMs, Kernels, Similarity Metrics, Deep Auto-encoders, Marginalised Stack Denoising Auto-encoders, t-SNE dimensionality reduction and density based clustering.

*Media:* Synoptica's blog article on me <https://goo.gl/WyeVpt>

### University of Surrey

SURREY, UNITED KINGDOM

#### Research Fellow

*Oct '15 – Nov '17*

My research focused on modelling biomedical changes over time with a view to forming actionable predictions using electronic medical records. I have been involved in developing machine learning algorithms to classify clinical time series trends and automatically identify acute events, along with algorithms to identify and automatically correct errors in medical records. As the biomedical measurements used are irregularly sampled, I also developed a method of re-sampling time series in order to enable them to be used when training a classifier. This has involved working closely with a core set of computational and clinical researchers, but also forming ad-hoc working relationships with an extended group of external clinical advisers and contributors.

*Programming & Tools:* Matlab

*Methods :* Gaussian Process Regression, Probabilistic Broken-sticks Modelling, Bayesian Modelling, Auto-encoders and t-SNE Clustering.

*Awards:* I was recognised with the **CogX 2018 award**: <https://goo.gl/QHnDpA>

## Lokal House Oy

HELSINKI, FINLAND

### Data Scientist

Jan '13 – Dec '13

My work involved crawling real estate websites to gather data and later converting the raw data into insights using natural language processing and advanced machine learning. The software I developed provided business analytics to potential investors.

*Programming & Tools* : Python, Scrapy, Django framework, BeautifulSoup, Lxml and MySQL.

*Methods:* Multilinear Regression Analysis, Gaussian Process Regression, Multiple Kernel Learning for Fusing Heterogeneous variables and Exponential Smoothing Techniques.

## Aalto University

ESPOO, FINLAND

### Research Assistant | Dept. of Marine Technology

Jun '12 – Dec '12

I was responsible for developing a software that extracts causal relations from the natural text present in the accident investigation reports of marine industry.

*Programming & Tools* : Python, Matlab and Natural Language Tool Kit (NLTK).

*Methods:* Modelling PoS tags and contextual grammar, Support Vector Machines (SVM) and Bayesian Classifiers.

### Research Assistant | Dept. of Energy Technology

Jun '11 – May '12

My work involved extracting meaningful representation from videos recordings of turbulent jets and sprays.

*Programming & Tools* : Matlab.

*Methods:* Proper Orthogonal Decomposition (POD), Image Analysis, Dynamic Mode Decomposition (DMD).

### Research Assistant | Information & Computer Science Lab

Oct '10 – May '11

My work involved clustering documents using different similarity metrics over different dimensionality reduction methods.

*Programming & Tools* : Java and Matlab

*Methods:* Principal Component Analysis (PCA), Singular Value Decomposition (SVD) and K-means clustering.

## Rofous Software Ltd

HYDERABAD, INDIA

### GIS Specialist

Mar '10 – Aug '10

I was appointed at Google India Pvt Ltd and was working at Google's office at Hyderabad on developing and testing the quality of labelled road markers on the Google maps.

*Programming & Tools* : Google's GIS tool and Google Earth.

## Medecode Solutions

HYDERABAD, INDIA

### Web Programmer & Web Developer

May '09 – Feb '10

I was developing web portals for several companies using Joomla content management system (CMS). I was also responsible for programming custom templates and modules within the Joomla CMS system.

*Programming & Tools* : PHP and Joomla CMS.

---

## Education

### University of Surrey

SURREY, UNITED KINGDOM

### Doctor of Philosophy (Ph.D.) in Computer Science

Jan 2014 – Nov 2016

I advanced the method of Dynamic Mode Decomposition for predicting, modelling and forecasting univariate time series. My thesis also introduced the method of DMD for motion correction and functional segmentation of anatomical structures in dynamical medical image sequences.

*Thesis:* Dynamic Mode Decomposition for Computer Vision and Signal Processing.

*Media:* During my PhD, I was recognised with three awards including the "Excellent Oral Presentation" from IEEE ICIVC <https://goo.gl/pQaZU5>, <https://goo.gl/PBJm8w> and "Best Research Potential" and "Outstanding Service" awards <https://goo.gl/nbZWor>

Aalto University

ESPOO, FINLAND

**Master of Science in Technology (MSc.(Tech.))** (by research)

Sep 2010 – Aug 2013

I graduated Masters' program with a major in machine learning and minor in computational fluid dynamics. My thesis focused on machine learning for natural language processing with an application to extract causal relations pertaining to accidents from the accident investigation reports.

*Thesis:* Mining Causal Relations from Maritime Accident Investigation Reports.

Jawaharlal Nehru Technological University

HYDERABAD, INDIA

**Bachelor of Technology (B. Tech)** (in Computer Science and Engineering)

Sep 2005 – May 2009

My dissertation focused around machine learning using neural networks to predict hand written characters as well as clustering high dimensional data.

*Thesis:* Machine Learning for Handwritten Character Recognition and Document Clustering.

---

## Skills

**Technical specialities:** Requirement gathering, algorithmic design, analysis and implementation.

I love Python, Matlab and  $\LaTeX$ .

Solid knowledge of web and cloud technologies: HTML+CSS+Jquery and AWS.

Linux administration skills: bash, Apache.

Database skills: MySQL.

Machine learning skills: Keras, Tensorflow, Scikit learn and Apache Solr.

**Machine/Deep learning theory:** I have strong skills in advanced machine/deep learning, multiple kernel learning, Neural nets, Bayesian methods, probabilistic methods, statistical signal processing, Gaussian process regression, support vector machines and high dimensional data clustering using t-stochastic neighbourhood embedding and self organisation maps. *I can write programmable code from mathematical expressions published in scientific articles.*

---

## Extra Curricular

Playing volleyball, running short distances, travelling and organising social events.

## Awards & Scholarships

2018	Outstanding Contribution to AI-Postdoc Research	CogX London 2018
2016	Best Presentation Award	IEEE ICIVC 2016
2015	Best Research Potential Award	PhD Conference
2015	Outstanding Service Prize	PhD Conference
2013	Full Studentship Grant	PhD in Computer Science
2013	Invention Award	Aalto Center for Entrepreneurship (ACE)
2012	Scholarship	Merenkulun säätiö
2012	Scholarship	Kotka Maritime Research Center
2011	Scholarship	Wärtsilä Finland Oy
2008	1 <sup>st</sup> /120 in Invention CIENCIA 2K8	Malware Doctor v1.1

## Strengths

I am always interested and motivated in work involving multi disciplinary fields. I have the ability to work, independently as well as also in a team. Furthermore, I have strong skills in mathematics.

## References

Available on request.

## Popular Scientific Publications

Please refer to my *Google scholar profile* for the complete list of articles. Current citation count: 297 (August 17, 2018)

1. **S Tirunagari**, N Poh, D Windridge, A Iorliam, N Suki, A Ho, *Detection of Face Spoofing Using Visual Dynamics*, Information Forensics and Security, IEEE Transactions on, 10 (4), 762 - 777, 2015.
2. **S Tirunagari**, K Wells, N Poh, M Bober, I Gordon, and D Windridge. *Movement Correction in DCE-MR Images through Windowed and Reconstruction Dynamic Mode Decomposition*. Machine Vision and Applications (2017) 28: 393. doi:10.1007/s00138-017-0835-5.
3. N Poh, **S Tirunagari**, S Bull, N Cole, S de Lusignan. *Probabilistic Broken-Stick Model: A Regression Algorithm for Irregularly Sampled Data with Application to eGFR*. Journal of biomedical informatics, 76, pp.69-77.
4. Vuorinen, V., J. Yu, **S Tirunagari**, O. Kaario, M. Larmi, C. Duwig, and B. J. Boersma. Large-eddy simulation of highly underexpanded transient gas jets, in Physics of Fluids 25, no. 1 (2013): 016101-016101.
5. **S Tirunagari**, N Poh, M Bober, D Windridge. *Can DMD Obtain a Scene Background in Color?* Image, Vision and Computing (ICIVC), International Conference on, 46-50, 2016. **(Best Presentation Award, ICIVC 2016)**.
6. **S Tirunagari**, N Poh, M Bober, D Windridge. *Windowed DMD as a Microtexture Descriptor for Finger Vein Counter-Spoofing in Biometrics*, Information Forensics and Security (WIFS), 2015 IEEE International Workshop on, 2015. **(Best Research Potential Award, PhDCompConf 2015)**.
7. **S Tirunagari**, S Bull and N Poh, Automatic Classification of Irregularly Sampled Time Series with Unequal Lengths: A Case Study on Estimated Glomerular Filtration Rate, IEEE International Workshop on Machine Learning for Signal Processing, Sept. 13-16, 2016, Salerno, Italy, 2016.
8. **S Tirunagari**, S Bull and N Poh. Visualisation of Survey Responses using Self-Organising Maps: A Case Study on Diabetes Self-care Factors. Computational Intelligence in eHealth and Health Care, (CICARE), IEEE Symposium on, 2016.
9. **S Tirunagari**, N Poh, K Aliabadi, D Windridge, D Cooke. Patient level analytics using self-organising maps: A case study on type-1 diabetes self-care survey responses, Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on, 2014.
10. S Kouchaki, **S Tirunagari**, A Tapinos and D L Robertson. Local Binary Patterns as a Feature Descriptor in Alignment-Free Visualisation of Metagenomic Data. Computational Intelligence and Data Mining (CIDM), Special Issue in Bioinformatics, IEEE Symposium on, 2016.
11. P Mari-Sanna, I Kivimäki, **S Tirunagari**, E Oja, and T Honkela. *Effect of dimensionality reduction on different distance measures in document clustering*. In International Conference on Neural Information Processing, pp. 167-176. Springer Berlin Heidelberg, 2011.