

## “1st BGSMath Data Science Workshop”

Barcelona, 22<sup>nd</sup> February 2017, Institut d'Estudis Catalans

### PROGRAM

---

The program includes presentations by BGSMath research groups from UAB, UB, UPC and UPF, a keynote lecture, presentations by companies and research centers followed by a roundtable on Data Science Innovation

9.00 Welcome

Marc Noy

Director, BGSMath

Marina Villegas

Directora, Agencia Estatal de Investigación

Francesc Subirada

Director General de Recerca

9.30 Jordi Vitrià (Universitat de Barcelona). "DataScience@UB: Going from theory to practice"

10.00 Piotr Zwiernik (Universitat Pompeu Fabra). "Sparsity in Gaussian totally positive distributions"

10.30 Coffee break

11.15 Ricard Gavaldà, Argimiro Arratia (Universitat Politècnica de Catalunya). "Research and applications at the LARCA (UPC) research group"

11.45 Joan del Castillo (Universitat Autònoma de Barcelona). "Advanced statistical modeling"

12.15 **Keynote Lecture**

Joan Bruna, Courant Institute, New York University

"Mathematics of Deep Learning: questions, conjectures, and some answers"

13.00 Lunch / Poster Session

14.30 Daniel Villatoro (Vodafone - Academia). "How to do evil with Data"

15.00 Cédric Notredame (Centre de Regulació Genòmica). "Harnessing numerical instability in high throughput sequence analysis pipelines using Nextflow"

15.30 Coffee break

16.00 David Torrents (Centro Nacional de Supercomputación / ICREA). "Biological data, more complex than big"

16.30 Roundtable Data Science innovation

José Antonio Rodríguez Serrano (BBVA), Pau Agulló (Kernel Analytics),  
Marc Torrent (Eurecat), Angel Faus (vLex)

17.15 Concluding remarks

Joan Bruna, Courant Institute, New York University  
Mathematics of Deep Learning: questions, conjectures, and some answers

**Abstract:**

In the age of ever increasing data and computational resources, the mathematical question of how to extract meaningful information from generic high-dimensional data remains a cursed one. However, deep learning architectures appear to be breaking this curse of dimensionality across many tasks in computer vision and speech recognition, using convolutional neural networks (cnn) that are trained by stochastic gradient descent.

This remarkable empirical success raises a number of puzzling questions spanning several statistical and mathematical fields. Why are deep convolutional architectures succeeding where other computer vision models didn't? Why is stochastic gradient descent so efficient for training such architectures despite the non-convexity? How can the successful inductive biases of cnn's be generalized to other signal domains?

In this talk, we shall connect these questions through a story of geometry, invariance and large deviations. Along the way, we will present a few conjectures and some answers, as well as open research directions at the interface of deep learning and mathematics.

Acevedo, Andrea	
<b>Automatic recognition system of peripheral blood cell images using deep features</b>	<b>1</b>
Alf3rez, Santiago	
<b>Automatic Classification of Normal, Reactive Lymphocytes, Abnormal Lymphoid Cells and Blast Cells.</b>	<b>2</b>
Arag3n, Pablo - G3mez, Vicenç	
<b>Measuring Platform Effects on Online Democracy</b>	<b>3</b>
Arenas, Concepcion	
<b>Supervised classification using a distance-depth function</b>	<b>4</b>
Cabaña, Alejandra	
<b>Permutation tests for the two sample problem for functional data</b>	<b>5</b>
Artiles, Luis	
<b>State space modeling strategies for censored time series a retail application</b>	<b>6</b>
Cobo L3pez, Sergio	
<b>Statistical Inference in Social Dilemmas</b>	<b>7</b>
Diaz Guilera, Albert	
<b>Quantification of network structural dissimilarities</b>	<b>8</b>
Ettamimi, Sara	
<b>Phylogenetic identification and characterization of microbial diversity in Moroccan rivers using a metagenomics approach</b>	<b>9</b>
Floden, Evan	
<b>Reproducible in-silico omics analyses across clouds and clusters with Nextflow</b>	<b>10</b>
Folguera Blasco, N3ria	
<b>Robustness of bistable gene regulatory circuits under epigenetic regulation: noise-enabled bifurcations in optimal transition path theory</b>	<b>11</b>
Guillen, Montserrat - Bolanc3, Catalina	
<b>Insurance Analytics: Predicting Probability of Customer Churn in Insurance</b>	<b>12</b>
Navas, V3ctor	
<b>Avalanches and force drops in displacement-driven compression of porous glasses</b>	<b>13</b>
Blazhievskaya, Irina	
<b>Properties of estimators of response functions in two-dimensional systems</b>	<b>14</b>
S3nchez Pla, Alex	
<b>Integrative Analysis to Select Genes Regulated by Methylation in a Cancer Colon Study</b>	<b>15</b>

Serra, Isabel	
<b>Deviation from power law of the global seismic moment distribution</b>	<b>16</b>
Vitrià, Jordi	
<b>Evaluating uncertainty scores for deep regression networks in financial short time series forecasting</b>	<b>17</b>
Zaiats, Vladimir	
<b>Asymptotically efficient statistical estimation in partially observed systems revisited</b>	<b>18</b>