

Daily Schedule [ Symposium 368 - ML in Astronomy: Possibilities and Pitfalls] 2-4 Aug 2022 (IAU GA, Busan, Rep. of Korea)


Session		IAUS368		Plenaries in Auditorium				
Date		2-Aug		Other talks in #205				
Time (KST, GMT+9)	Program		Category	Speakers			Type of Participation	Remarks
	Program	Slot		Name	Abs. no.	Title		
08:15-09:45	Registration		Registration					1st day of GA - Registration
09:45-10:30	Morning e-Poster							
10:30-12:00	Morning Oral Session	S368-1	Opening Remarks	SOC				15 minutes
			Invited	Sara Webb	3253	ML tutorial for the broader community	Remote	1:15 hours
12:00-13:30				Lunch				
13:30-15:00	Afternoon Oral Session 1	S368-2	Invited	Guillermo Cabrera	3009	Classic Machine Learning vs Deep Learning: when, why and how?	In-room	1:30 hours
15:00-15:15				Break				
15:15-16:45	Afternoon Oral Session 2	S368-3	Invited	Michelle Lochner	1545	Machine Learning in Astronomy	Remote	45 Minutes
			Invited	Benjamin L'Huillier, Mark Huertas Company (+), Giorgos Varnardos, Joshua Speagle (+)		Panel discussion: Broader ML Topics	In-room	45 Minutes
16:45-17:30	Afternoon e-Poster							

Date		3-Aug						
Time (KST, GMT+9)	Program		Category	Speakers			Type of Participation	Remarks
	Program	Slot		Name	Abs. no.	Title		
08:15-09:45	Plenary Lecture		Plenary	George Djorgovski	1791	Machine Learning in Astronomy: From the Star-Galaxy Separation to a Collaborative Human-AI Discovery	Remote	45 Minutes
			Plenary	Ofer Lahav	2189	Deep Learning in Astronomy: Trends and Challenges	In-room	45 Minutes
09:45-10:30	Morning e-Poster							
10:30-12:00	Morning Oral Session	S368-4	Invited	Renee Hlozek	1870	Existing data sets for machine learning in Astronomy	In-room	45 Minutes
			Invited	Nikhita Madhanpall, Mike Walmsley, Shay Zucker		Panel discussion: Methodology for fusion of large datasets	In-room	45 Minutes
12:00-13:30				Lunch				
13:30-15:00	Afternoon Oral Session 1	S368-5	Contributed	David Parkinson	555	Detecting complex sources in large surveys using an apparent complexity measure	In-room	15 Minutes
			Contributed	Dennis Crake	2021	In Search of the Peculiar: An Unsupervised Approach to Anomaly Detection in the Transient Universe.	In-room	15 Minutes
			Contributed	Didier Fraix-Burnet	1003	Unsupervised classification: a necessary step for Deep Learning?	In-room	15 Minutes
			Contributed	Gordian Edenhofer	1049	Iterative Grid Refinement: Approximate Gaussian Processes for Billions of Parameters	In-room	15 Minutes
			Contributed	Jeroen Audaeraert	1830	Unraveling the physical mechanisms of pulsating stars through a multimodal and multidisciplinary machine learning approach	In-room	15 Minutes
			Contributed	Annalisa Pillepich	2178	ERGO-ML: Extracting Reality from Galaxy Observables with Machine Learning	In-room	15 Minutes
15:00-15:15				Break				
15:15-16:45	Afternoon Oral Session 2	S368-6	Invited	Yuan-Sen Ting, Sungwook Hong, Lilliane Nakazono, Guillermo Cabrera		Discussion: Practical Problem Solving - including interpretability	In-room	90 Minutes
16:45-17:30	Afternoon e-Poster							

Date		4-Aug						
Time (KST, GMT+9)	Program		Category	Speakers			Type of Participation	Remarks
	Program	Slot		Name	Abs. no.	Title		
08:15-09:45	Plenary Lecture		Plenary					Symposium 369
09:45-10:30	Morning e-Poster							
10:30-12:00	Morning Oral Session	S368-7	Invited	Eric Ford	1546	Enhancing Exoplanet Surveys via Physics-informed Machine Learning	Remote	45 Minutes
			Invited	Samaya Nisanke, Jae-Hun Jung, Melissa Lopez, Rafael Martinez Galarza		Panel Discussion: GW/MMA	In-room	45 Minutes
12:00-13:30				Lunch				
13:30-15:00	Afternoon Oral Session 1	S368-8	Invited	Ivy Wong	927	A review of current tools for outreach & education	In-room	45 Minutes
			Contributed	Melissa Lopez	1744	Simulating Transient Noise Bursts in LIGO with Generative Adversarial Networks	In-room	15 Minutes
			Contributed	Mike Walmsley	1150	Galaxy Zoo: Practical Methods for Large-Scale Learning	In-room	15 Minutes
			Contributed	Joshua Speagle	707	Incorporating Errors in Machine Learning Methods	Remote	15 Minutes
15:00-15:15				Break				
15:15-16:45	Afternoon Oral Session 2	S368-9	Contributed	Raquel Ruiz Valenca	2501	Comparing machine learning and deep learning models to estimate quasar photometric redshifts	In-room	15 Minutes
			Contributed	Steffani Grondin	1203	Searching for the extra-tidal stars of Galactic globular clusters with high-dimensional clustering analysis	In-room	15 Minutes
			Contributed	Vishal Upendran	515	Accelerating astronomy workflow with deep learning and interpretable AI	In-room	15 Minutes
			Contributed	Yuan-Sen Ting	503	Quantifying non-Gaussianity with mathematical insights from machine learning	In-room	15 Minutes
			Summary	SOC		Meeting Summary and Next Steps	In-room	15 Minutes
			Closing Remarks	SOC			In-room	15 Minutes
16:45-17:30	Afternoon e-Poster							

\*+ Remote

## S 368 at IAU GA 2022: Machine Learning - Possibilities and Pitfalls - Report

Meeting Identification Number: Symposium 368  
Meeting Title: Machine Learning - Possibilities and Pitfalls  
Coordinating Division: B  
Location (city, country): Busan, Republic of Korea  
Dates of meeting: 2-4 Aug 2022  
Number of participants: 450+ interested; 100+ in-person; many online  
Total Amount of IAU Grant funds received (in euros): 2 registration discounts + 2 waivers  
Number of IAU Grant recipients: 4  
List of represented countries: Same as IAU GA  
Anticipated number of separate papers in the proceedings: 90  
Report submitted by: Ashish Mahabal  
Date and place: 12 Oct 2022, Pasadena, CA, USA  
Signature of SOC Chairperson: 

Symposium 368 of the IAU GA in Busan on “Machine Learning - Possibilities and Pitfalls” was held from Aug 2 to 4 2022. The theme revolved around machine learning in astronomy, in particular the potential it has created to solve many open challenges, but also the possible problems that can be encountered when applying these often black-boxy techniques without proper care.

### Speakers:

The eight invited speakers are listed below and include four men and four women. In addition, there were 13 contributed talks given by nine males, and four females. There were also four panel discussions during which 15 total panelists shared the stage. The complete program is appended. For want of time another ~70 abstracts had to be converted to e-talks and e-posters. They were available throughout the meeting.

Name	Abs no.	Title
Eric Ford	1546	Enhancing Exoplanet Surveys via Physics-informed Machine Learning
George Djorgovski	1791	Machine Learning in Astronomy: From the Star-Galaxy Separation to a Collaborative Human-AI Discovery
Guillermo Cabrera	3009	Classic Machine Learning vs Deep Learning: when, why and how?
Michelle Lochner	1545	Machine Learning in Astronomy
O. Ivy Wong	927	A review of current tools for outreach & education
Ofer Lahav	2189	Deep Learning in Astronomy: Trends and Challenges
Renee Hlozek	1870	Existing datasets for machine learning in Astronomy
Sara Webb	3253	ML tutorial for the broader community

**Panelists:** Benjamin L’Huillier, Mark Huertas Company (+), Giorgos Vernardos, Joshua Speagle (+)

Nikhita Madhanpall, Mike Walmsley, Shay Zucker, Yuan-Sen Ting, Sungwook Hong, Lilliane Nakazono, Guillermo Cabrera, Samaya Nissanke, Jae- Hun Jung, Melissa Lopez, Rafael Martinez Galarza (+ == Remote)

**Chair persons:** Ashish Mahabal, Kai Polsterer, David Parkinson, Vanessa McBride, Arman Shafieloo, Renee Hlozek, Didier Fraix-Burnet, Annalisa Pillepich

### **Highlights in brief:**

The symposium started with two detailed tutorials “ML tutorial for the broader community” by Sara Webb and “Classic Machine Learning vs Deep Learning: when, why and how?” by Guillermo Cabrera. These provided a gentle introduction to those who were not steeped in the details of machine learning, providing them just the impetus to be able to follow the other talks later. For those already well-versed they served as a quick revision.

The plenary talks addressed the trends and challenges in deep learning (Ofer Lahav) addressing how deep learning contrasts with shallow learning, and the challenges of explainability and interpretability as more deep learning techniques are embraced; and the path to collaborative Human-AI learning (George Djorgovski) addressing how with modern ML/AI methods for the first time we are starting to discover patterns from data driven approaches and how this leads to areas not investigated before.

Other invited talks included reviews of various areas. Michelle Lochner set up things with an overview of Machine Learning in Astronomy, Renee Hlozek reviewed the existing and forthcoming large and diverse datasets that form staple inputs to ML algorithms, Eric Ford explained how physics-informed machine learning is starting to play an important role in exoplanet characterization, and Ivy Wong detailed citizen science aspects for creating labeled datasets and for outreach and education.

The discussion sessions covered various aspects of the advances in machine learning, and the possible unintended misuse thereof. The first one discussed the breadth of ML applications in astronomy highlighting the fact that for many problems classical ML techniques may be the first refuge, but also that recent improvements in techniques mean that one does not need huge datasets to do machine learning. Another discussed the important topic about combining diverse datasets, in particular archival with real-time. A third took on practical problem solving including interpretability. Finally, the fourth discussion handled more recent trends in multi-messenger astronomy. Open discussions saw strong participation of students and postdocs..

The contributed talks - chosen from a large pool of abstracts - provided more extensive coverage of the field, touching upon more surveys, outliers and anomalies, unsupervised methods, incorporation of errors, non-Gaussianity, Generative Adversarial Networks etc. The speakers included not only senior members but also undergraduate and graduate students. What these talks may not have covered was certainly covered by the e-talks and e-posters providing for a well-rounded symposium with science topics from our neighborhood to cosmology.

### **Additional thoughts and comments:**

We received overwhelming support with close to 450 astronomers wanting to attend the symposium in person, and about 100 submitting abstracts for presentations. We tried to select representative presentations from across the spectrum, with the talks covering theory, simulations, applications, interpretations and on datasets from the Solar System to Galactic to extragalactic astronomy to cosmology.

There were many more deserving abstracts that had to be converted to e-talks and e-posters. We ensured that there was ample time for discussion through various panels on topics like Gravitational Wave/Multi-Messenger Astronomy, broader ML, and fusion of large datasets, in which all attendees were able to participate.

The challenges that ML is facing include the lack of interpretability and explainability. On another level, not many techniques allow for proper uncertainty quantification. While we did not expect the symposium to solve these big problems, we did manage to grow awareness amongst a larger set of users and practitioners of the issues, moving the needle towards better ML practices.

The broad hope and expectation was that attendees gain exposure to the breadth of available datasets and techniques, and the expertise on display during the symposium allows them to step out of their comfort zone to take on bigger problems in a safe manner. In particular, we hope to see more population studies incorporating more publicly available datasets (including transfer learning across datasets) rather than specialized studies involving smaller private and proprietary datasets, and we believe the symposium has helped us move in that direction.

Since the pandemic had not yet abated, and the corresponding health safety measures were still in place, we likely lost many interested attendees (even from the 15 SOC members, only five could attend in person). Also, there were many sessions in parallel further dividing the attendees. Yet, throughout the symposium over 100 attendees were present. Plus many online, though we did not have a way of knowing how many. There were a few complaints from those online about not being able to follow everything in the room - this is the unfortunate reality of hybrid meetings.

There were ample opportunities for discussions which the attendees took advantage of. There was dedicated time for poster viewing, but since these were e-posters it is likely that they got less visibility during the meeting. But having the posters available to view for an extended period may make up for that.

Overall the symposium set things up well to invite novices into the world of machine learning with datasets large and small, and provided the experts with more fodder to explore new problems, datasets, and techniques while taking steps towards avoiding the associated pitfalls.