

Machine Learning in Space



esa

Atılım Güneş Baydin
gunes@robots.ox.ac.uk

University of Oxford Space Tech Conference
7 Nov 2023



Outline

- About
- Frontier Development Lab (NASA, ESA, Oxford)
- Machine learning and space
 - Case I: constellations and collision avoidance
 - Case II: thermospheric density estimation
- Machine learning onboard spacecraft
 - Case III: flood detection from space, ML in Low Earth Orbit
 - Case IV: measuring molecular complexity, ML onboard robotic missions
- Community

Atılım Güneş Baydin

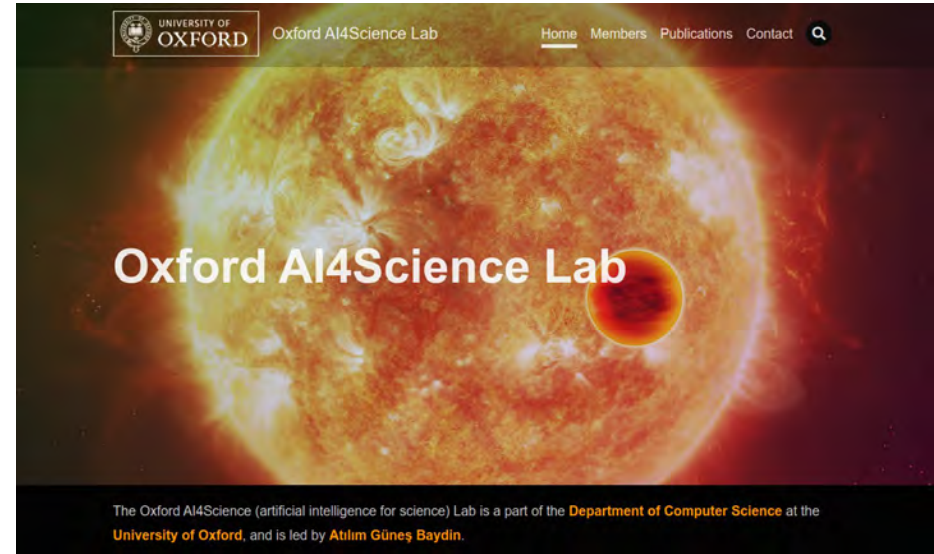
Departmental Lecturer in Machine Learning, Department of Computer Science



Oxford AI4Science (Artificial Intelligence for Science) Lab

<https://oxai4science.github.io>

- Specializing in **probabilistic machine learning and scientific discovery**
- Working with experts in high-energy physics, heliophysics, astrobiology, Earth science, space safety and other disciplines
- Solve challenging problems through application and development of AI methods



Collaborators at:



Frontier Development Lab
NASA, ESA, Oxford

FDL 2022

EUROPE



Google Cloud



SCAN^oAI



AIRBUS



CATAPULT



Frontier Development Lab

<https://fdl.ai>

- A research accelerator for state-of-the-art ML and space sciences
- Two main versions
 - NASA Ames & SETI Institute (FDL US)
 - ESA & University of Oxford (FDL Europe)
- Access to compute provided by industry (Google, Intel, Nvidia and others)
- Teams of
 - PhD students / postdocs (two machine learning, two domain science)
 - supervising faculty



Machine learning and space

CASE I

Altitude (Km) :

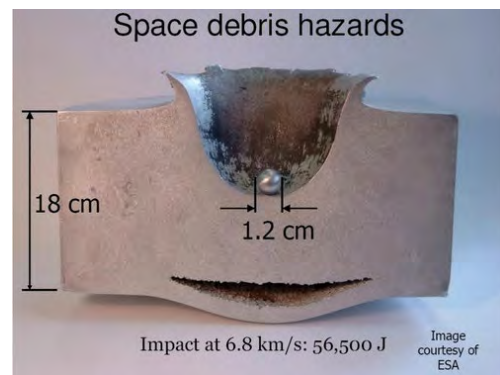
19 121

ML for spacecraft collision avoidance



Low Earth orbit (LEO) has millions of uncontrolled objects (~1 cm) at ~28,000 km/h orbital speeds

- **Kessler Syndrome:** chain reaction of collisions can pollute the orbit and render it inaccessible
- Danger to satellites, scientific missions, future human access
- Large constellations, e.g., SpaceX StarLink (40k), OneWeb and rapidly increasing commercial space activities contribute to the problem



Acciarini, Pinto, Metz, Boufelja, Kaczmarek, Merz, Martinez-Heras, Letizia, Bridges, **Baydin** "Kessler: a Machine Learning Library for Space Collision Avoidance" **8th European Conference on Space Debris 2021**

Pinto, Acciarini, Metz, Boufelja, Kaczmarek, Merz, Martinez-Heras, Letizia, Bridges, **Baydin** "Towards Automated Satellite Conjunction Management with Bayesian Deep Learning" **AI Earth Sci Workshop, NeurIPS 2020**

Acciarini, Pinto, Metz, Boufelja, Kaczmarek, Merz, Martinez-Heras, Letizia, Bridges, **Baydin** "Spacecraft Collision Risk Assessment with Probabilistic Programming" **ML4PS workshop, NeurIPS 2020**

ML for spacecraft collision avoidance



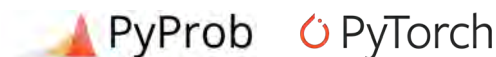
Worked with ESA Space Debris Office to automate risk assessment and maneuver decisions

- Working with conjunction data messages (CDMs)
- Probabilistic programming with physical orbit simulators that generate scenarios and synthetic CDMs
- Bayesian deep learning to predict event evolution as a sequence of CDMs



<https://github.com/kesslerlib>

Named after Donald J. Kessler



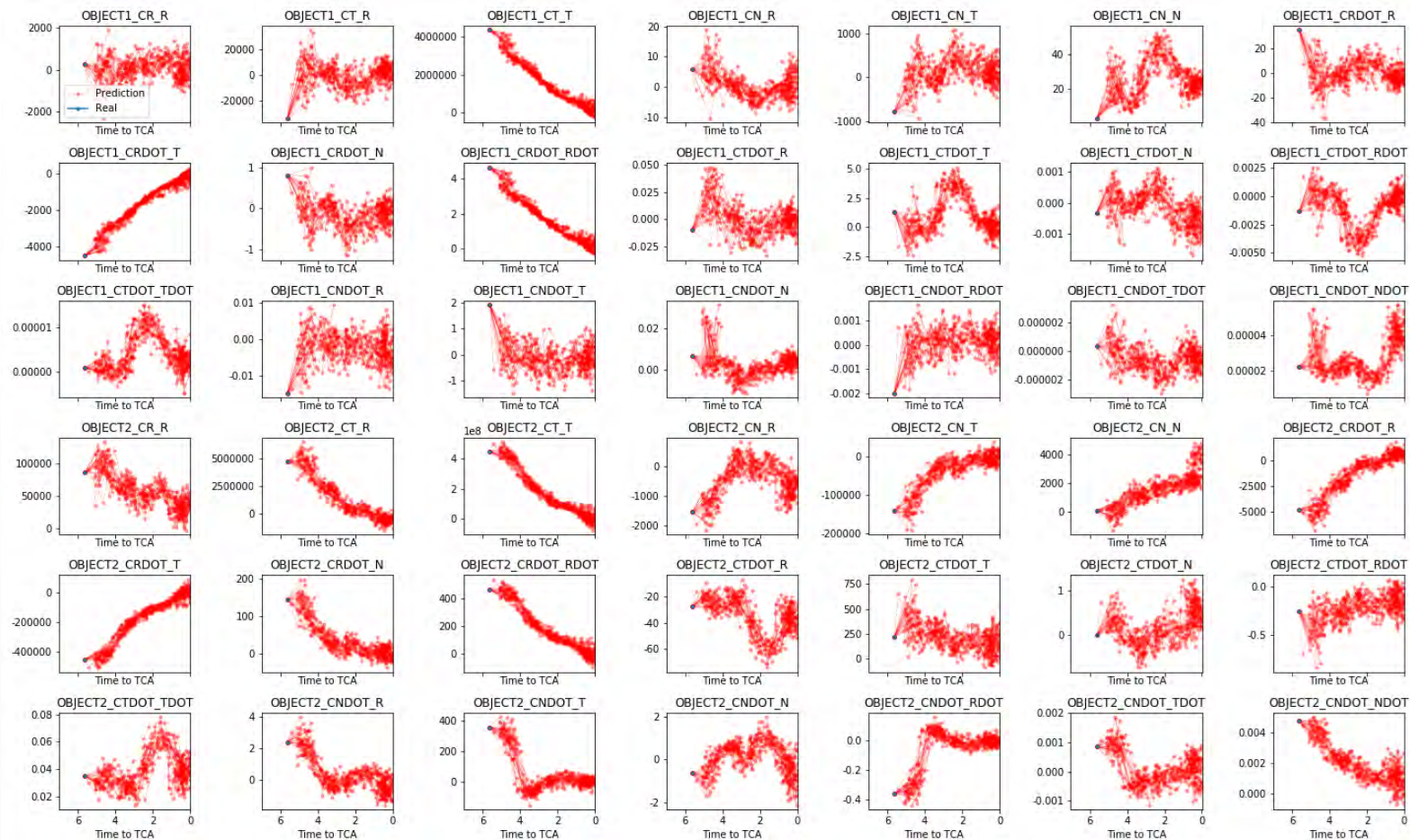
Funding: ESA ESRIIN / PhiLab, Google Cloud

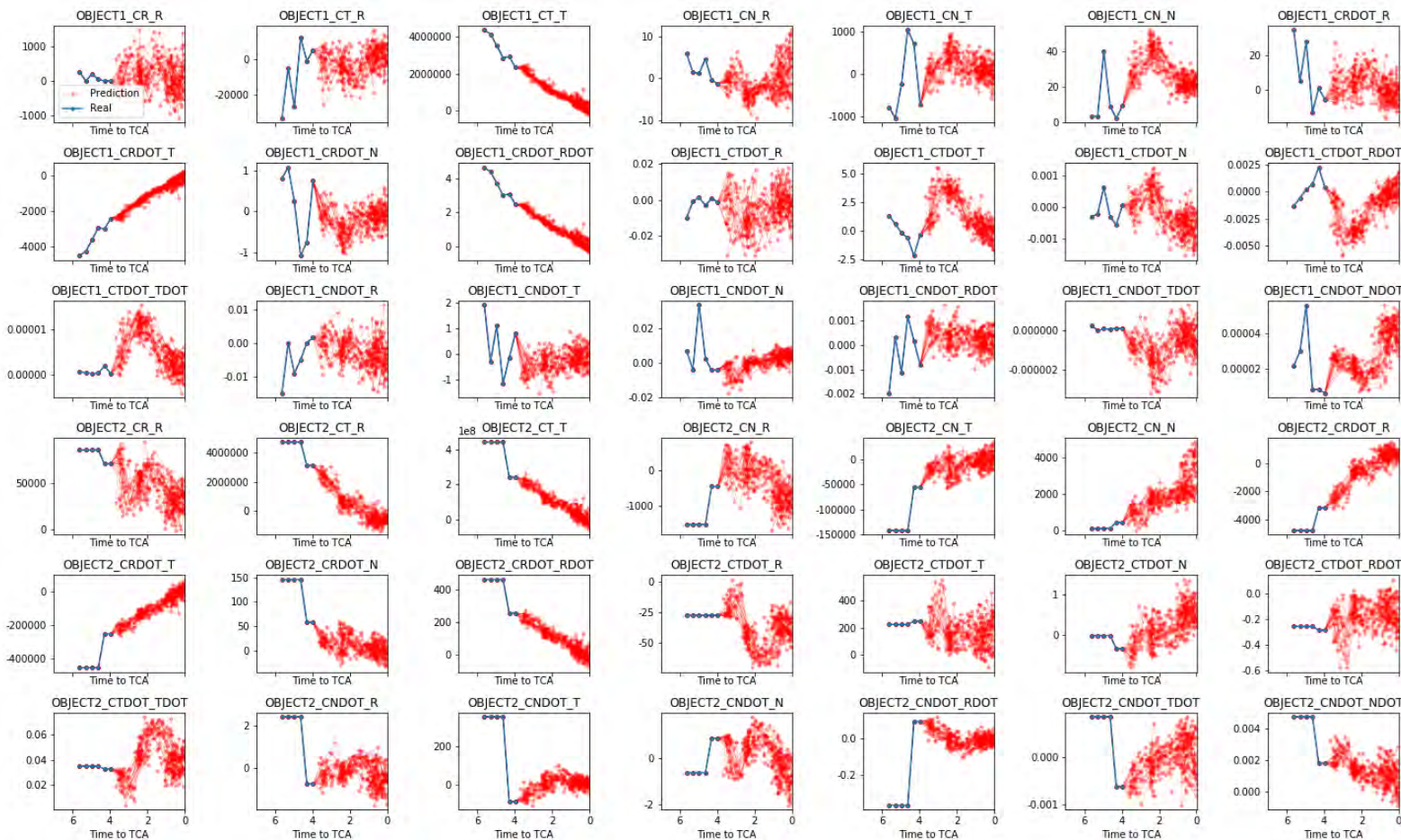
Funding: UK Space Agency (£250k, with Surrey Space Centre and Cranfield)

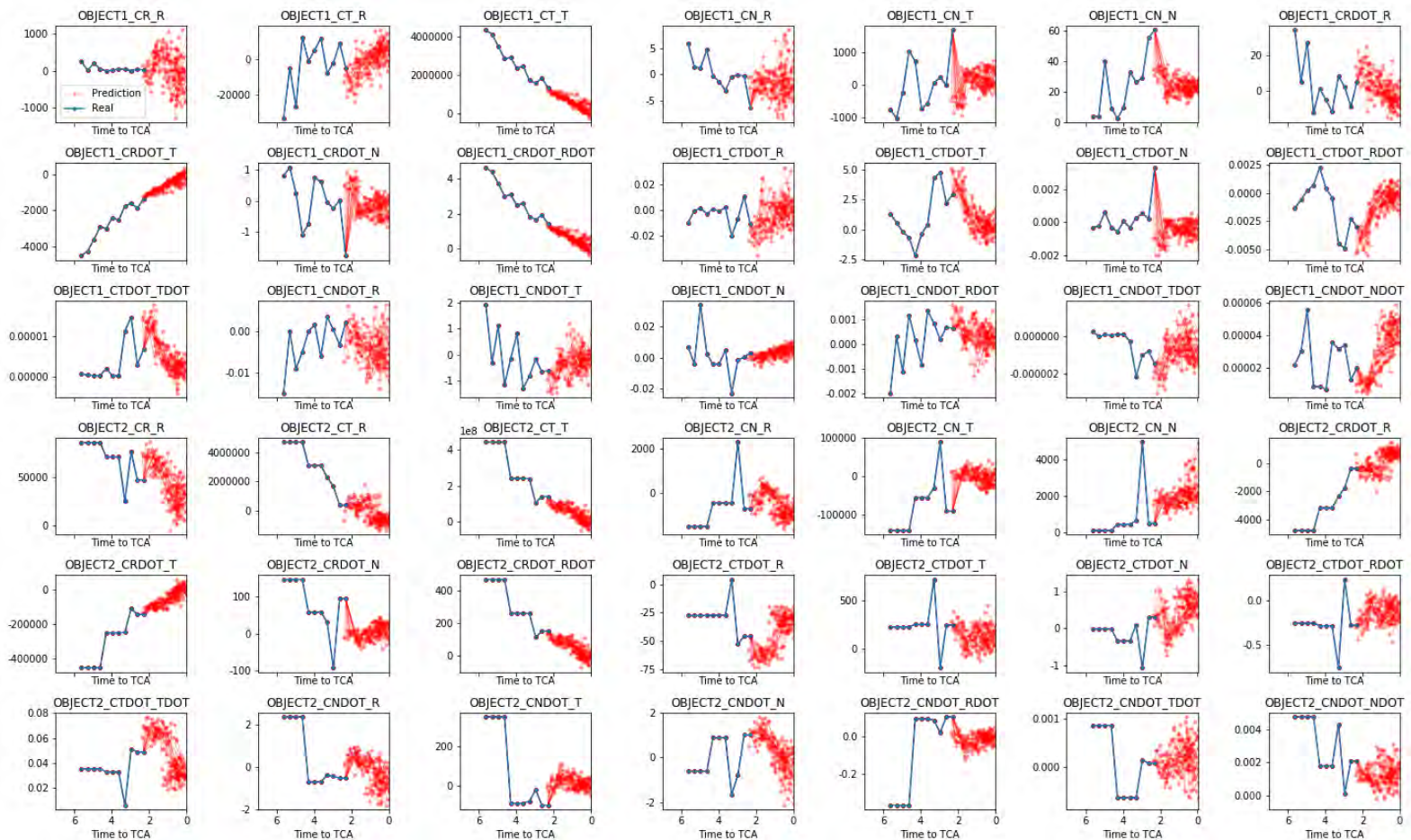
Acciarini, Pinto, Metz, Boufelja, Kaczmarek, Merz, Martinez-Heras, Letizia, Bridges, **Baydin** "Kessler: a Machine Learning Library for Space Collision Avoidance" **8th European Conference on Space Debris 2021**

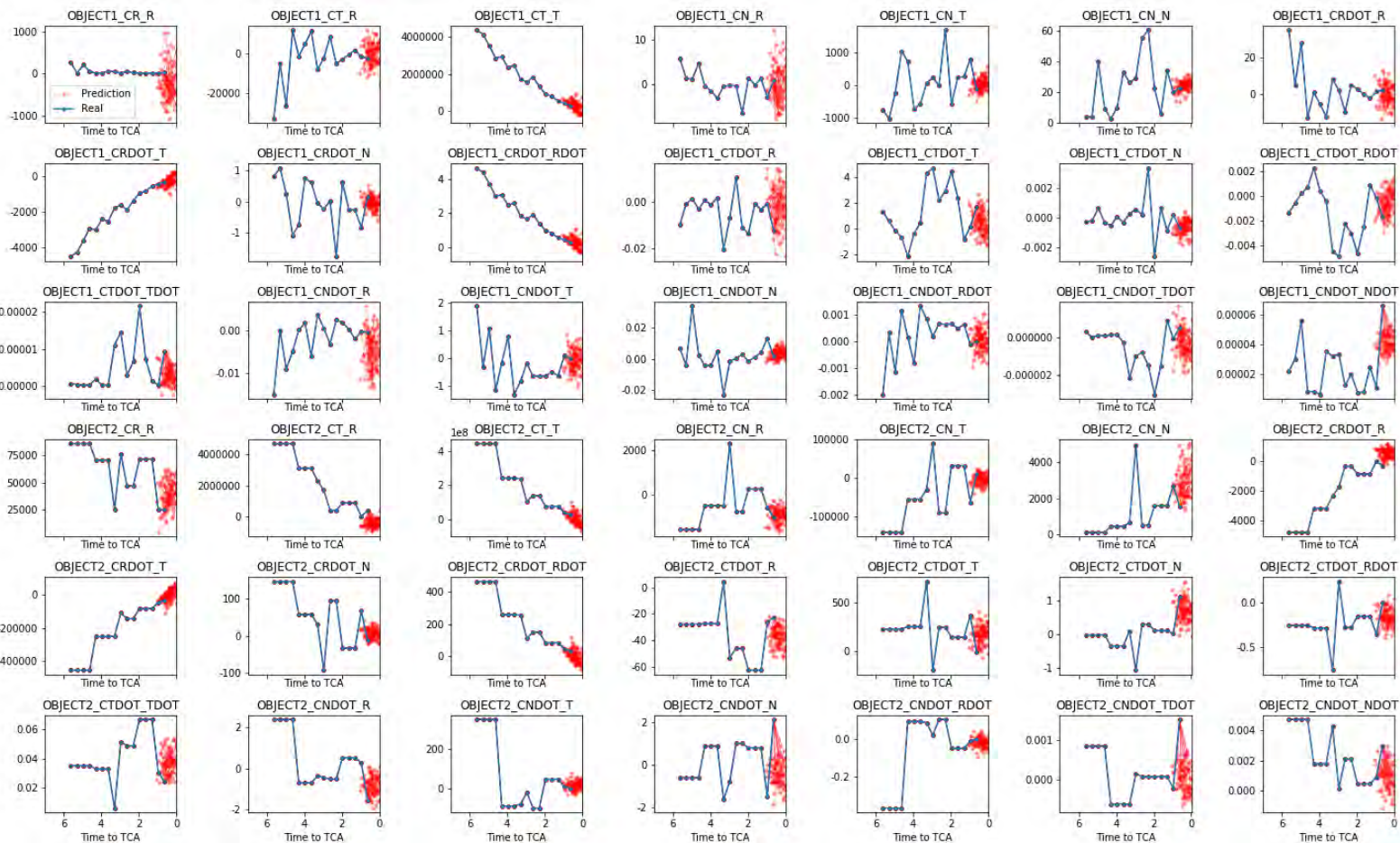
Pinto, Acciarini, Metz, Boufelja, Kaczmarek, Merz, Martinez-Heras, Letizia, Bridges, **Baydin** "Towards Automated Satellite Conjunction Management with Bayesian Deep Learning" **AI Earth Sci Workshop, NeurIPS 2020**

Acciarini, Pinto, Metz, Boufelja, Kaczmarek, Merz, Martinez-Heras, Letizia, Bridges, **Baydin** "Spacecraft Collision Risk Assessment with Probabilistic Programming" **ML4PS workshop, NeurIPS 2020**









A photograph taken from space showing the Earth's horizon. The atmosphere is a vibrant blue, and a bright green aurora borealis is visible in the upper atmosphere. The text "CASE II" is overlaid in the center.

CASE II

Thermospheric density estimation using machine learning



Thermospheric density models are **a key ingredient in mission planning and operations** in Earth orbit

- Orbital motion is affected by gravity and non-gravitational perturbations (drag)
- **Thermospheric density** is a major component of drag, affected by solar activity
- Current models are not accurate, computationally expensive, based on **proxies** for of solar activity (solar irradiance $F10.7$, geomagnetic indices ap , Kp)

Work requested by NASA SMD to develop benchmarks of thermospheric density estimation using ML with direct solar data from in-orbit instruments (NASA SDO)

Funding: NASA SMD Heliophysics Division



SpaceX loses 40 satellites to geomagnetic storm a day after launch

9 February

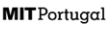


GETTY IMAGES

This Falcon 9 rocket launch on 3 February carried 49 Starlink satellites, most of which were caught by the storm

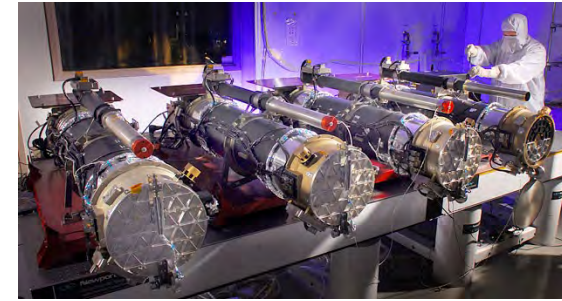
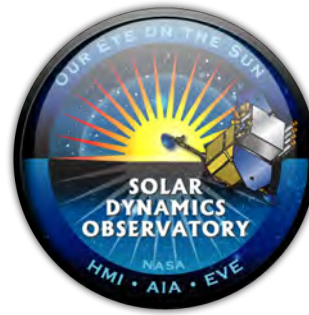
SpaceX has lost dozens of satellites after they were hit by a geomagnetic storm a day after launch, causing them to fall from orbit and burn up.

Such solar "storms" are caused by powerful explosions on the sun's surface.

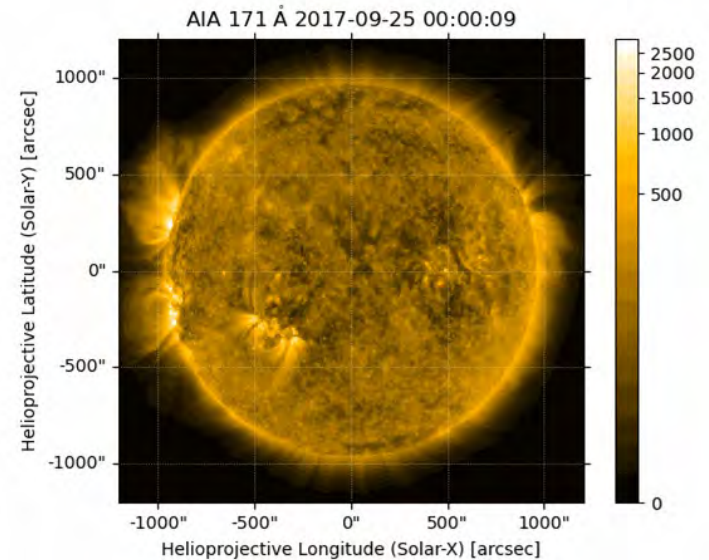


Thermospheric Density Estimation

- State-of-the-art thermospheric models governed by **on-ground irradiance proxies/geomagnetic indices**
 - $F10.7$ → Solar irradiance
 - ap, Kp → Geomagnetic field
- Solar data available from **in-orbit instruments such as the Solar Dynamics Observatory (SDO)**
- Curated solar data available (SDOML dataset)
- Machine learning techniques can leverage data to
 - Improve and forecast solar proxies
 - Generate surrogate atmospheric models
 - Drive knowledge discovery for other planets/moons



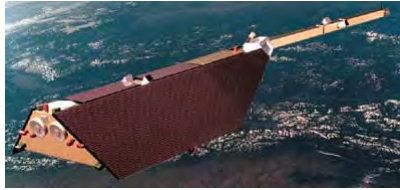
SDO AIA instrument



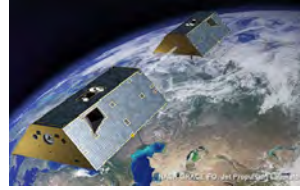
Thermospheric Density Estimation: Combining Data from Multiple Missions

Trained with ground truth data from

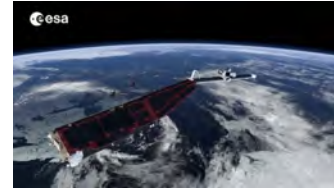
- CHAMP
- GRACE
- Swarm A - B
- GOCE



CHAMP (GFZ), 2001 - 2010



GRACE (NASA), 2002 - 2009

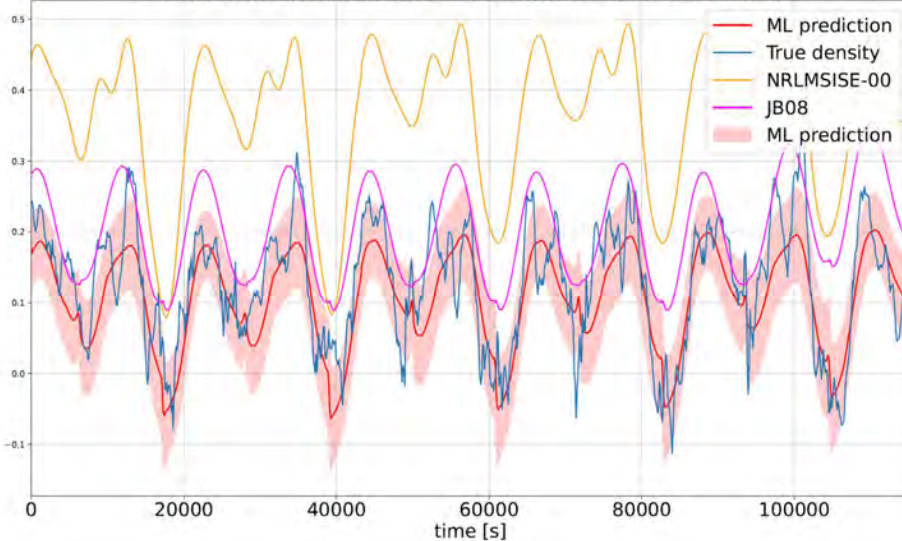


Swarm (ESA), 2013 - 2021

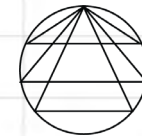
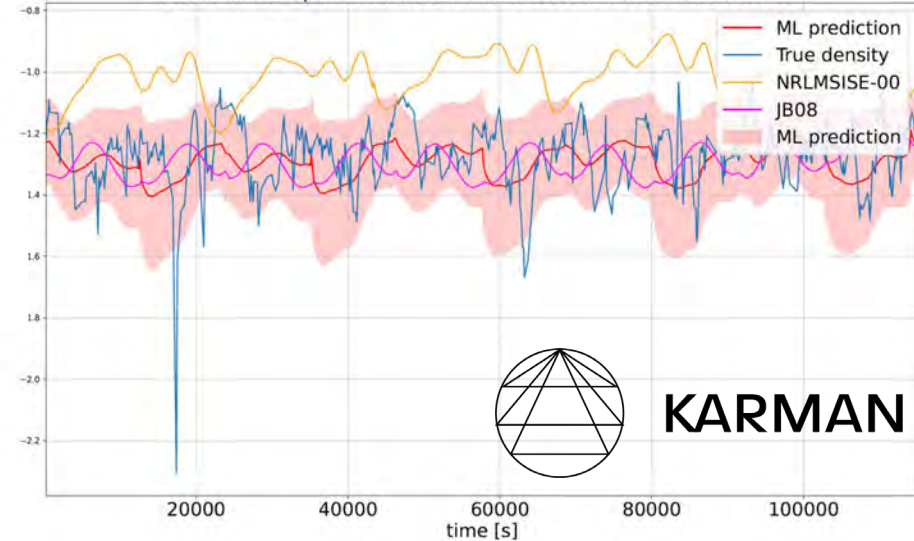


GOCE (ESA), 2010 - 2013

Champ: FFNN Experiment 2008-12-15 00:00:00 - 2008-12-15 08:00:00



Grace: FFNN Experiment 2008-12-15 00:00:00 - 2008-12-15 08:00:00



KARMAN



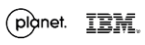
Google Cloud



intel.



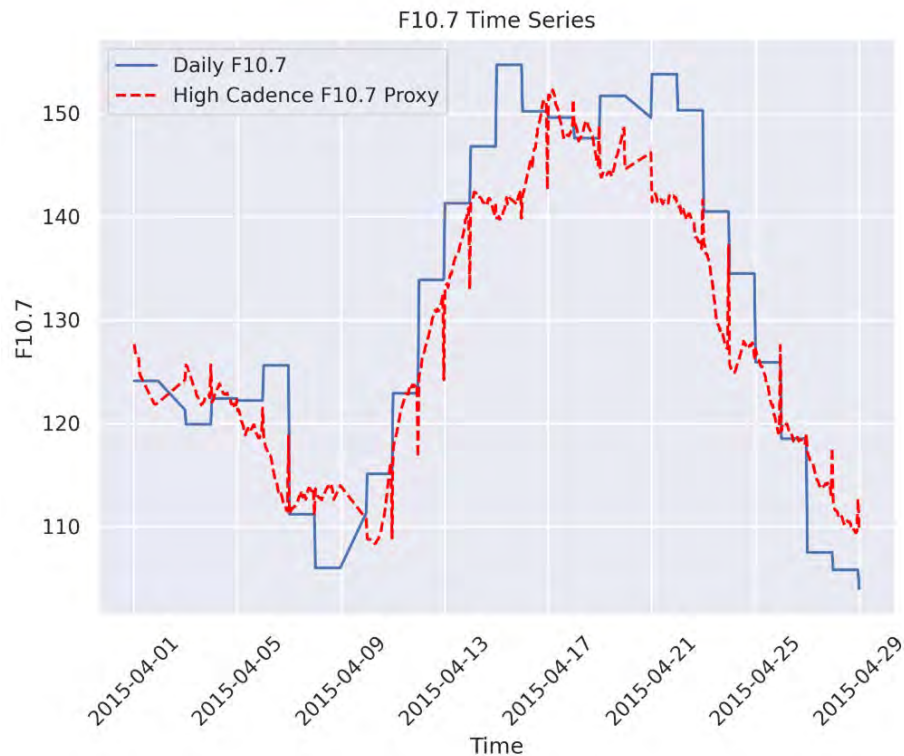
MIT Portugal





KARMAN

Key Result:
ML-informed replication of F10.7
at higher cadence.



Machine learning on board

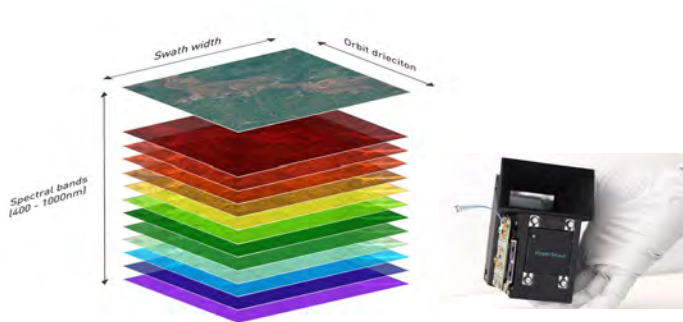
SpaceX Transporter 2 launch,
30 June 2021

CASE III

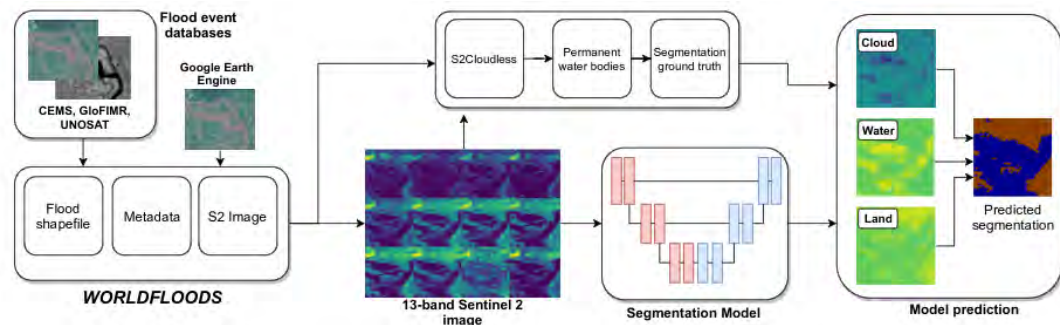
A photograph of a SpaceX Falcon Heavy rocket launching. The rocket is positioned vertically in the center of the frame, ascending into a sky filled with large, white, billowing clouds. A bright, intense plume of fire and white smoke trails behind the rocket, extending from its base to the top of the frame. Below the rocket, the launch pad structure is visible, including several tall, slender service towers. The foreground consists of a dense line of green trees and bushes. The overall scene is dramatic and captures the power of the launch.

Deploying ML in space

- CubeSats can provide affordable **disaster response at a fraction of cost of larger missions**, making it accessible to developing countries
- Perform **flood segmentation onboard satellite** to reduce downlinked data, using weight-quantized U-nets and convolutional neural networks



50-band hyperspectral camera



Mateo-Garcia, Veitch-Michaelis, Smith, Oprea, Schumann, Gal, **Baydin**, Backes "Towards Global Flood Mapping Onboard Low Cost Satellites with Machine Learning" **Scientific Reports 2021** (in press)

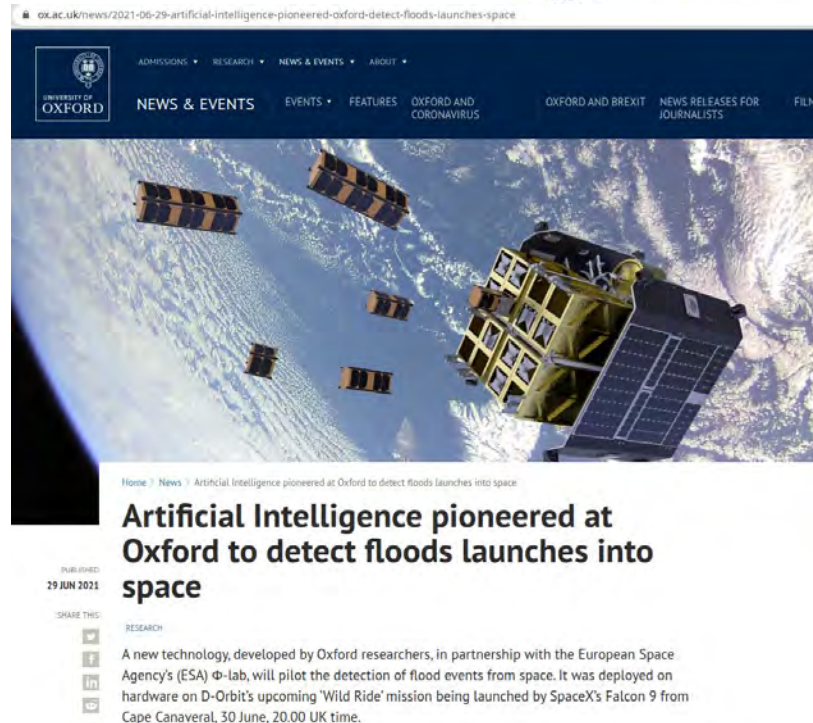
Mateo-Garcia, Oprea, Smith, Veitch-Michaelis, **Baydin**, Backes "Flood Detection On Low Cost Orbital Hardware." **AI for Humanitarian Assistance and Disaster Response Workshop, NeurIPS 2019**

Deploying ML in space

Worked with ESA Centre for Earth Observation (ESRIN)

- Originally targeted the PhiSat-1 (Sep 2020)
- **First step:** demonstrated feasibility on Intel hardware (Movidius ML accelerator) identical with the one on board satellite
- **Second step:** Pilot the detection of flood events from space on **D-Orbit's** Wild Ride mission
- Successfully **launched on SpaceX Transporter-2 Mission**, 30 June 2021

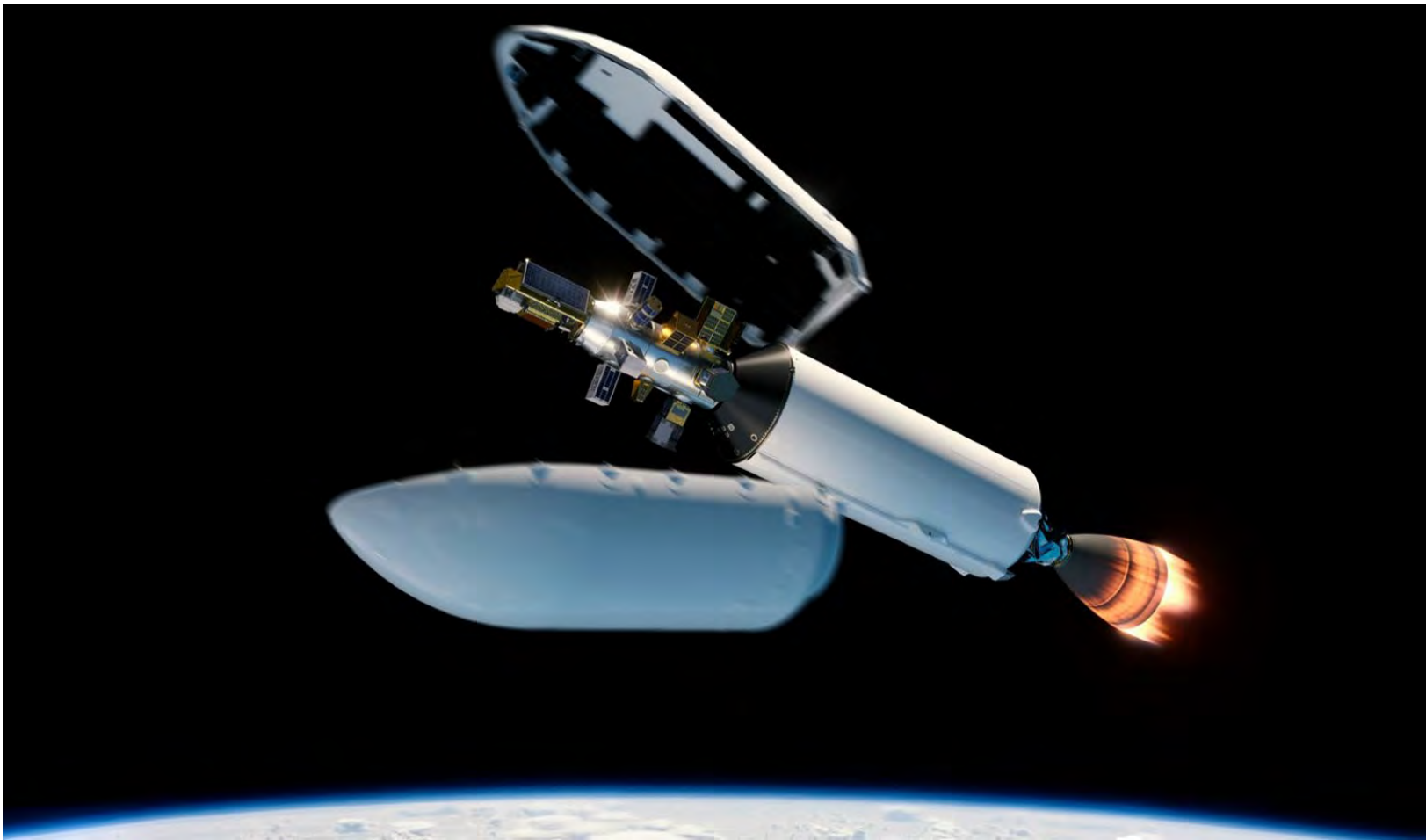
<https://www.ox.ac.uk/news/2021-06-29-artificial-intelligence-pioneered-oxford-detect-floods-launches-space>



Mateo-Garcia, Veitch-Michaelis, Smith, Oprea, Schumann, Gal, **Baydin**, Backes "Towards Global Flood Mapping Onboard Low Cost Satellites with Machine Learning" **Scientific Reports 2021**

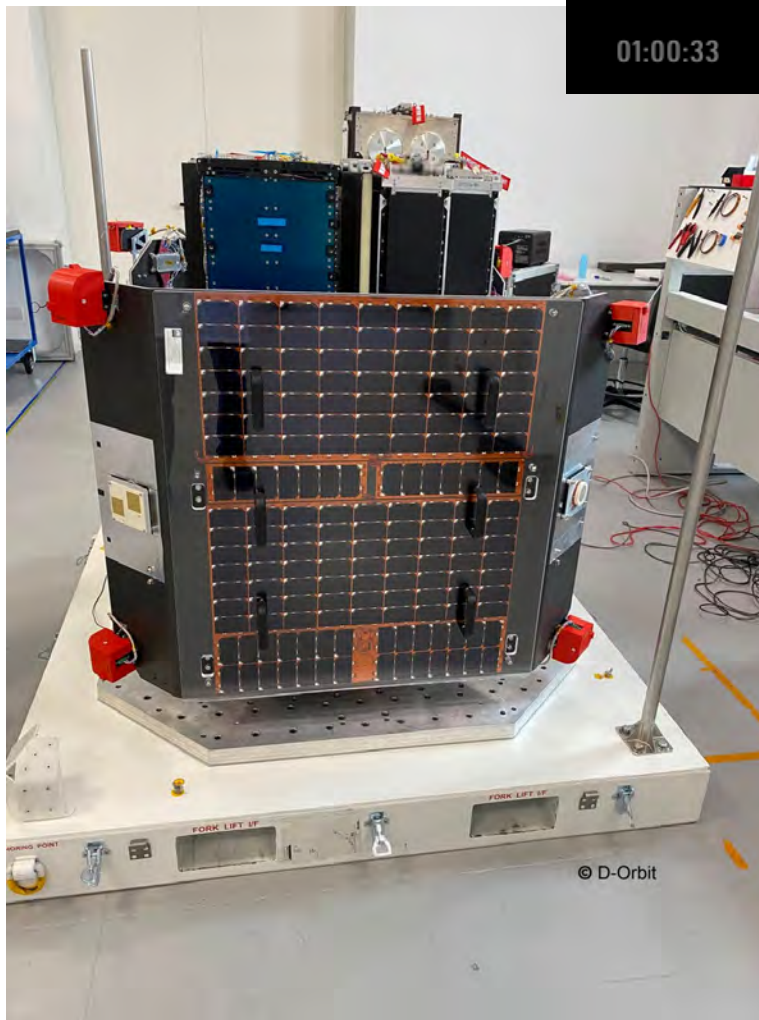
Mateo-Garcia, Oprea, Smith, Veitch-Michaelis, **Baydin**, Backes "Flood Detection On Low Cost Orbital Hardware." **AI for Humanitarian Assistance and Disaster Response Workshop, NeurIPS 2019**

“Transporter 2”





ION Satellite Carrier



01:00:33

D-Orbit's ION satellite carrier deploys





“Transporter 2”





“Transporter 2”

**SpaceX Falcon 9
rocket.**

**Space Launch
Complex 40
(SLC-40)**












FRONTIER DEVELOPMENT LAB EUROPE

WORLDFLOODS Dataset

-  150 floods
-  618 flood maps
-  235,000 patches (256x256 px)
-  303 GB

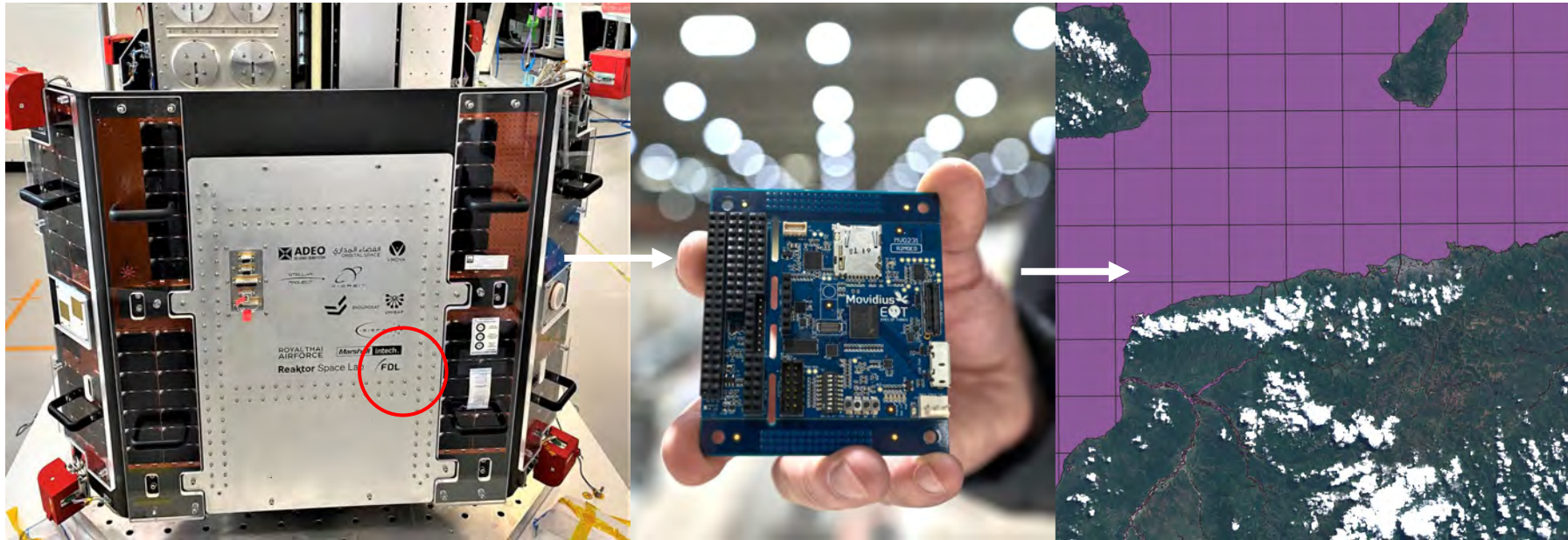
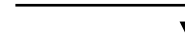


A world map showing the locations of floods. The map is color-coded with blue, green, and yellow dots indicating flood locations across various continents, including North America, Europe, and Asia.



Worldfloods + Space Cloud

Processing S2 imagery in orbit to obtain a flood segmentation.





Worldfloods

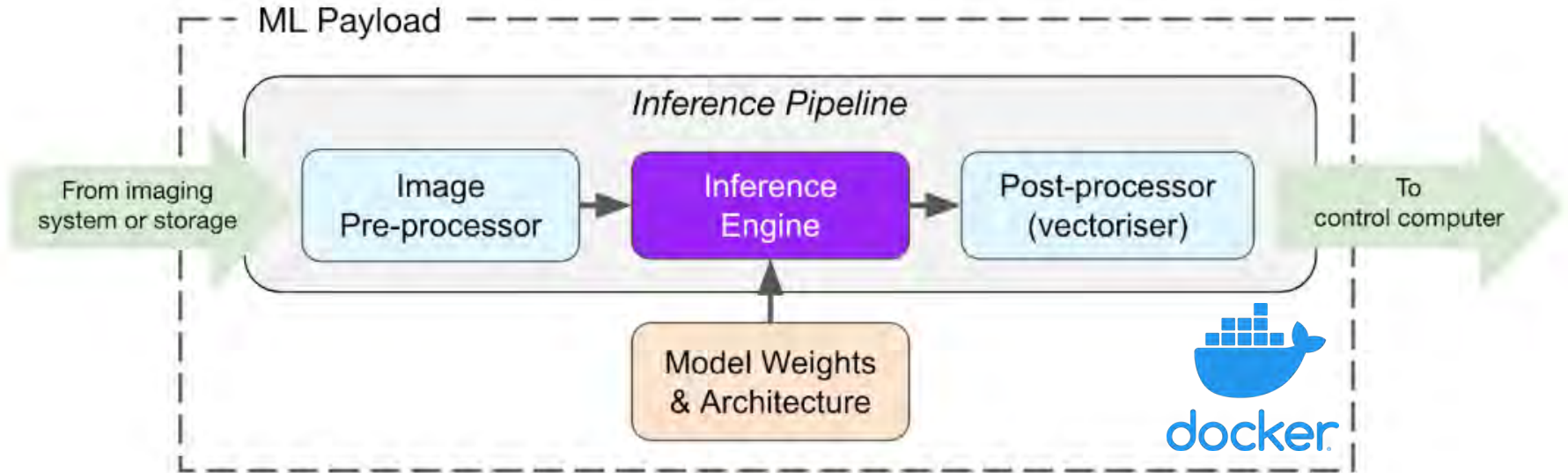




Worldfloods + Space Cloud



Dockerized Inference pipeline deployed in D-Orbit hardware



NASA DragonFly Mission to
Saturn/Titan

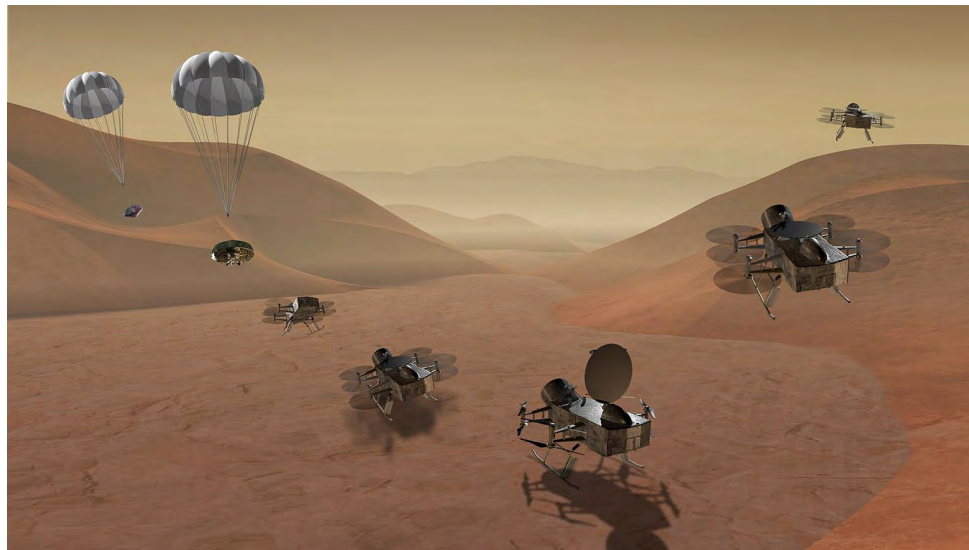
CASE IV

An aerial, high-angle view of a vast, arid desert landscape. The terrain is characterized by rolling sand dunes and a network of winding, eroded paths or dry riverbeds. The color palette is dominated by warm, earthy tones of tan, beige, and light brown, with some darker shadows cast by the dunes. The sky is a pale, hazy yellow, suggesting a dusty or hazy atmosphere. In the upper left corner, a small, dark, curved object is visible, likely a part of the spacecraft or camera lens.



ONBOARD DETECTION OF MOLECULAR BIOSIGNATURES

- Dragonfly is a NASA mission to Saturn/Titan (launch planned June 2027)
- Robotic rotorcraft (VTOL)
- Instruments
 - Dragonfly Mass Spectrometer (DraMS)
 - Dragonfly Gamma-Ray and Neutron Spectrometer (DraGNS)
 - Dragonfly Geophysics and Meteorology Package (DraGMet)



We are working on ML methods to show feasibility of running an onboard biosignature detection with data from DraMS

<https://www.nasa.gov/dragonfly>



Defining “life” is hard!

What if alien life looks nothing like anything we’ve ever seen on Earth?



Illustrations: flaticon.com

WHAT IS MOLECULAR COMPLEXITY | WAYS TO MEASURE IT

How big is it?

Da: Mass of atoms in the molecule

How diverse are its atomic ingredients?

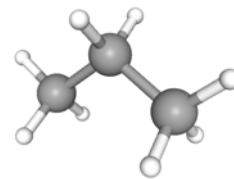
C_m: Number and nature of bonds in a molecule and asymmetry

What holds it together?

BHI: Number of bonds + diversity of atoms in that make up the molecule

How is it made?

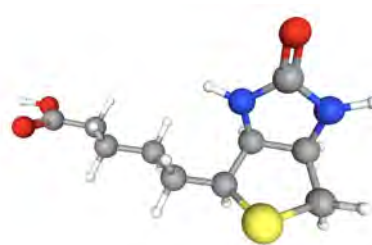
MA: Pathway to synthesize a given molecule



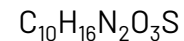
Propane



44.11 g/mol (Daltons)



Biotin



244.31 g/mol (Daltons)

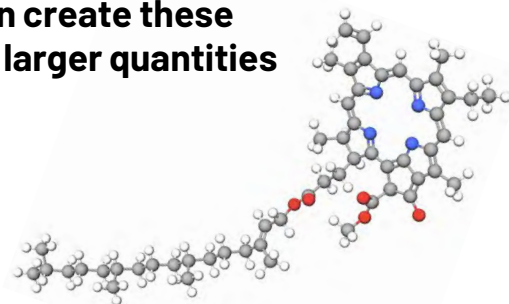
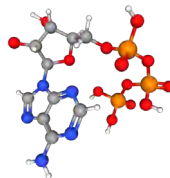
HYPOTHESIS¹ | "LIFE" IS LIKELY TO PRODUCE COMPLEXITY-IN-ABUNDANCE

Threshold?!

These molecules can be formed through abiotic processes



Only "life" can create these molecules in larger quantities



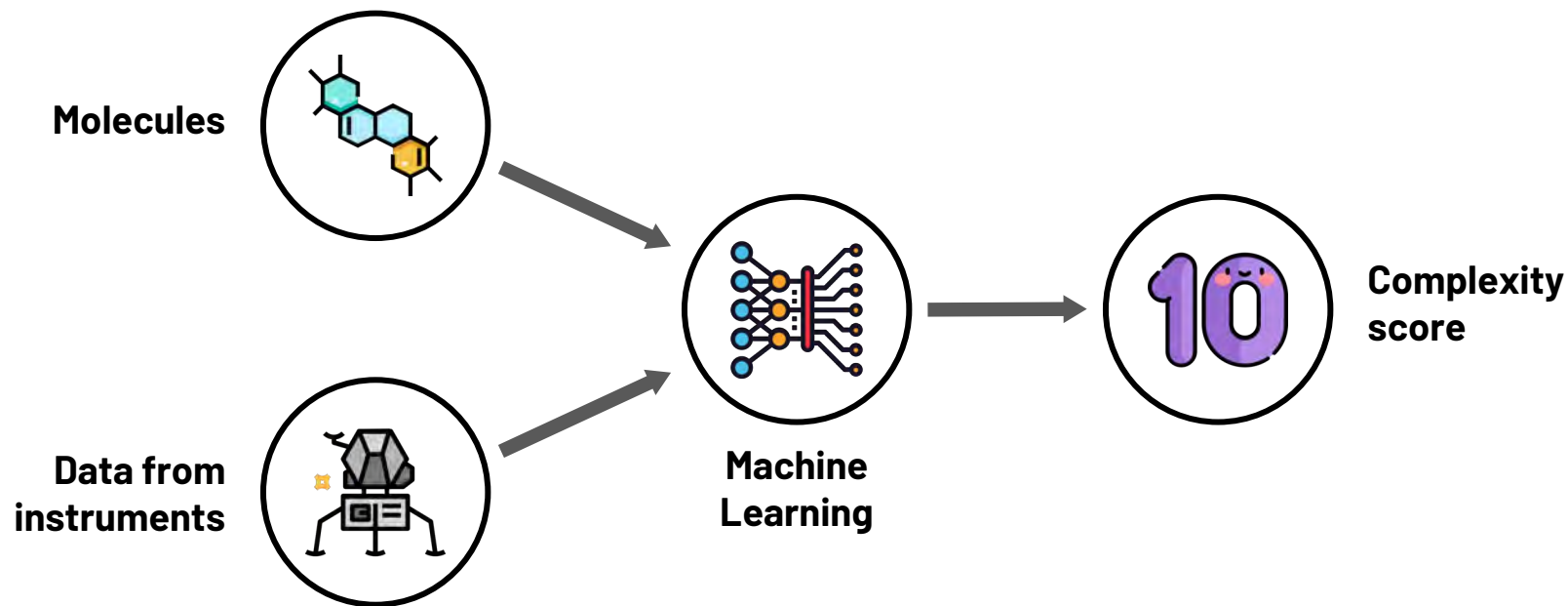
1

20

Increasing complexity scores →

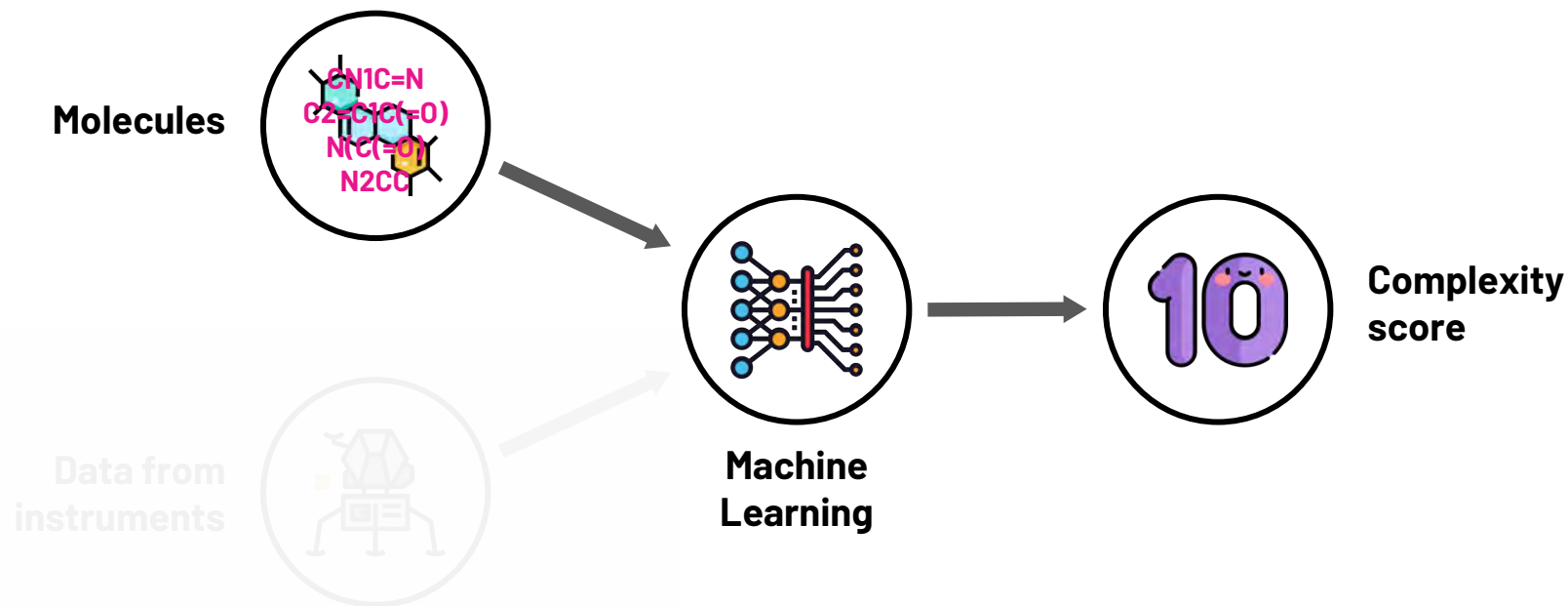
¹See, e.g., Marshall et al. (2017, 2021)

MACHINE LEARNING | OUR WORKFLOW



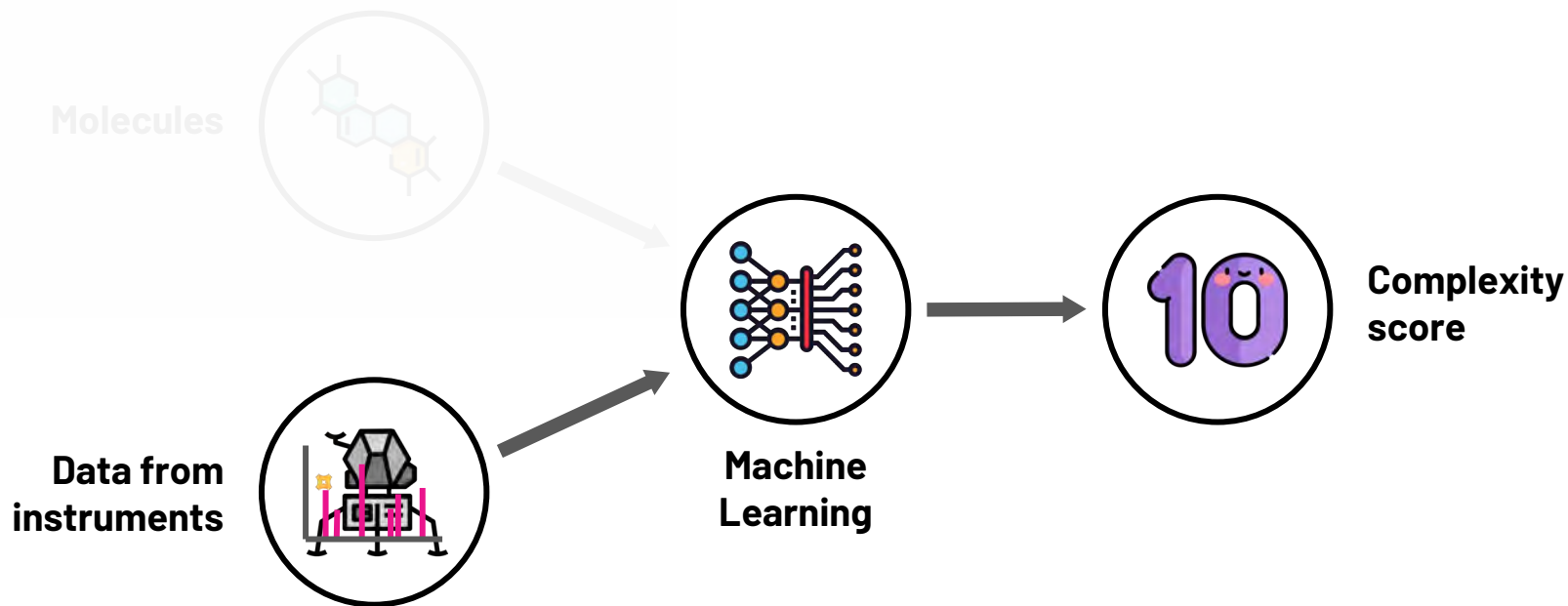
Illustrations: flaticon.com

MACHINE LEARNING | OUR WORKFLOW



Illustrations: flaticon.com

MACHINE LEARNING | OUR WORKFLOW



Illustrations: flaticon.com

DATASET | MOLECULAR DATA AGGREGATED FROM MANY SOURCES

NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

PubChem

ChEMBL

NIH NATIONAL CANCER INSTITUTE

Reaxys

Google Scholar

...

query
collect

complexipy

Our code
(and brains!)

combine
compute

150,000+
unique molecules

14,000+
mass spectra

100,000+
CPU node hours

DATASET | CURATING DATA MEANS LOTS OF “INVISIBLE” WORK!

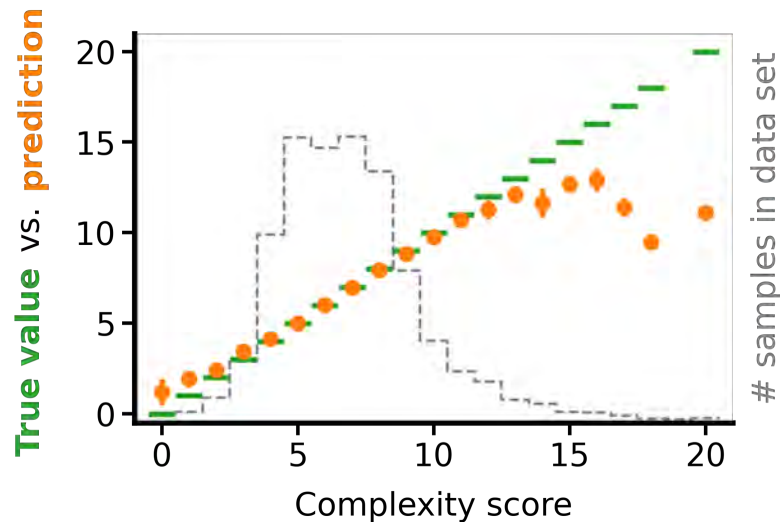
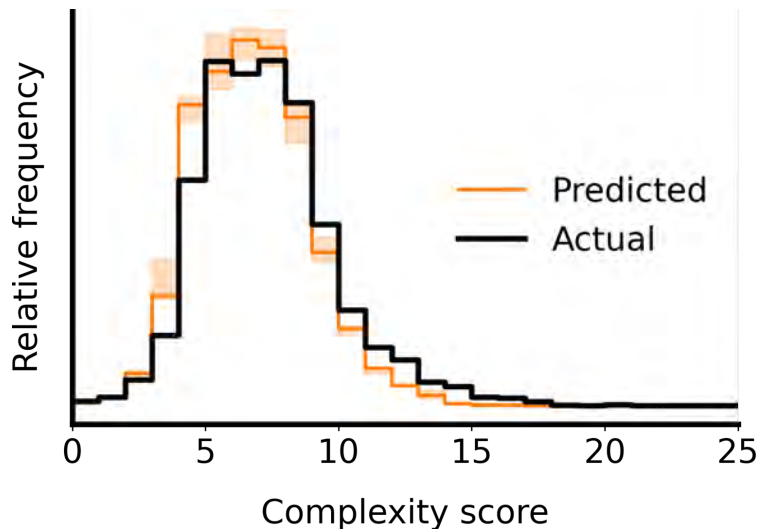
Machine Learning



Data Curation

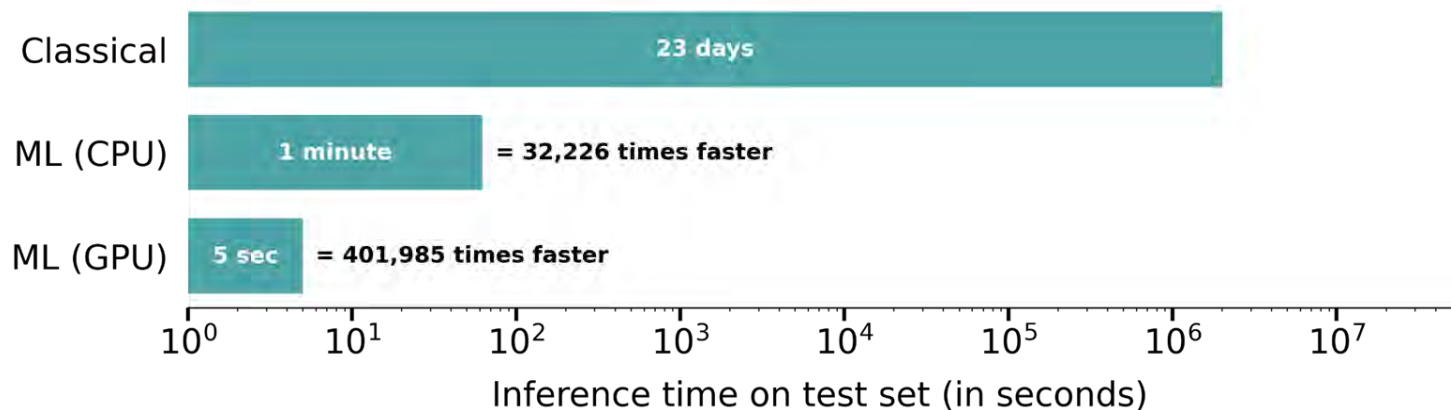
Illustration: freepik.com

MACHINE LEARNING | RESULTS: MASS-SPEC-TO-COMPLEXITY



MACHINE LEARNING | RESULTS: MOLECULE-TO-COMPLEXITY

Test set: ~4000 molecules, mean prediction error: **9.8%**



Thank you for listening

Questions?

Selected references

- [1] T. A. Le, A. G. Baydin, and F. Wood. Inference compilation and universal probabilistic programming. In International Conference on Artificial Intelligence and Statistics (AISTATS), 2017.
- [2] A. Munk, A. Ścibior, A. G. Baydin, A. Stewart, G. Fernlund, A. Poursartip, and F. Wood. Deep probabilistic surrogate networks for universal simulator approximation. In PROBPROG, 2020.
- [3] A. G. Baydin, L. Heinrich, W. Bhimji, L. Shao, S. Naderiparizi, A. Munk, J. Liu, B. Gram-Hansen, G. Louppe, L. Meadows, P. Torr, V. Lee, Prabhat, K. Cranmer, and F. Wood. Efficient probabilistic inference in the quest for physics beyond the standard model. In NeurIPS, 2019.
- [4] A. G. Baydin, L. Shao, W. Bhimji, L. Heinrich, L. F. Meadows, J. Liu, A. Munk, S. Naderiparizi, B. Gram-Hansen, G. Louppe, M. Ma, X. Zhao, P. Torr, V. Lee, K. Cranmer, Prabhat, and F. Wood. Etalumis: Bringing probabilistic programming to scientific simulators at scale. In SC19, 2019.
- [5] K. Cranmer, J. Brehmer, and G. Louppe. The frontier of simulation-based inference. Proceedings of the National Academy of Sciences, 117(48):30055–30062, 2020.
- [6] B. Gram-Hansen, C. Schroeder, P. H. Torr, Y. W. Teh, T. Rainforth, and A. G. Baydin. Hijacking malaria simulators with probabilistic programming. In ICML workshop on AI for Social Good, 2019.
- [7] B. Gram-Hansen, C. S. de Witt, R. Zinkov, S. Naderiparizi, A. Scibior, A. Munk, F. Wood, M. Ghadiri, P. Torr, Y. W. Teh, A. G. Baydin, and T. Rainforth. Efficient bayesian inference for nested simulators. In AABI, 2019.
- [8] B. Poduval, A. G. Baydin, and N. Schwadron. Studying solar energetic particles and their seed population using surrogate models. In MML for Space Sciences workshop, COSPAR, 2021.
- [9] G. Acciarini, F. Pinto, S. Metz, S. Boufelja, S. Kaczmarek, K. Merz, J. A. Martinez-Heras, F. Letizia, C. Bridges, and A. G. Baydin. Spacecraft collision risk assessment with probabilistic programming. In ML4PS (NeurIPS 2020), 2020.
- [10] F. Pinto, G. Acciarini, S. Metz, S. Boufelja, S. Kaczmarek, K. Merz, J. A. Martinez-Heras, F. Letizia, C. Bridges, and A. G. Baydin. Towards automated satellite conjunction management with bayesian deep learning. In AI for Earth Sciences Workshop (NeurIPS), 2020.
- [11] G. Acciarini, F. Pinto, S. Metz, S. Boufelja, S. Kaczmarek, K. Merz, J. A. Martinez-Heras, F. Letizia, C. Bridges, and A. G. Baydin. Kessler: a machine learning library for space collision avoidance. In 8th European Conference on Space Debris, 2021.
- [12] S. Shirobokov, V. Belavin, M. Kagan, A. Ustyuzhanin, and A. G. Baydin. Black-box optimization with local generative surrogates. In NeurIPS, 2020.
- [13] H. S. Behl, A. G. Baydin, R. Gal, P. H. S. Torr, and V. Vineet. Autosimulate: (quickly) learning synthetic data generation. In 16th European Conference on Computer Vision (ECCV), 2020.
- [14] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. Automatic differentiation in machine learning: a survey. Journal of Machine Learning Research (JMLR), 18(153):1–43, 2018.
- [15] A. G. Baydin, B. A. Pearlmutter, and J. M. Siskind. DiffSharp: An AD library for .net languages. In 7th International Conference on Algorithmic Differentiation, 2016.
- [16] A. G. Baydin, R. Cornish, D. M. Rubio, M. Schmidt, and F. Wood. Online learning rate adaptation with hypergradient descent. In ICLR, 2018.
- [17] H. Behl, A. G. Baydin, and P. H. Torr. Alpha maml: Adaptive model-agnostic meta-learning. In AutoML (ICML), 2019.
- [18] A. G. Baydin, K. Cranmer, M. Feickert, L. Gray, L. Heinrich, A. Held, A. Melo, M. Neubauer, J. Pearkes, N. Simpson, N. Smith, G. Stark, S. Thais, V. Vassilev, and G. Watts. Differentiable programming in high-energy physics. In Snowmass 2021 Letters of Interest (LOI), Division of Particles and Fields (DPF), American Physical Society, 2020.
- [19] A. G. Baydin, K. Cranmer, P. de Castro Manzano, C. Delaere, D. Derkach, J. Donini, T. Dorigo, A. Giammanco, J. Kieseler, L. Layer, G. Louppe, F. Ratnikov, G. C. Strong, M. Tosi, A. Ustyuzhanin, P. Vischia, and H. Yarar. Toward machine learning optimization of experimental design. Nuclear Physics News International (Submitted), 2020.
- [20] L. F. Guedes dos Santos, S. Bose, V. Salvatelli, B. Neuberger, M. Cheung, M. Janvier, M. Jin, Y. Gal, P. Boerner, and A. G. Baydin. Multi-channel auto-calibration for the atmospheric imaging assembly using machine learning. Astronomy & Astrophysics (in press), 2021.
- [21] A. D. Cobb, M. D. Himes, F. Soboczenski, S. Zorzan, M. D. O’Beirne, A. G. Baydin, Y. Gal, S. D. Domagal-Goldman, G. N. Arney, and D. Angerhausen. An ensemble of bayesian neural networks for exoplanetary atmospheric retrieval. The Astronomical Journal, 158(1), 2019.
- [22] C. Schroeder de Witt, B. Gram-Hansen, N. Nardelli, A. Gambardella, R. Zinkov, P. Dokania, N. Siddharth, A. B. Espinosa-Gonzalez, A. Darzi, P. Torr, and A. G. Baydin. Simulation-based inference for global health decisions. In ICML Workshop on Machine Learning for Global Health, Thirty-seventh International Conference on Machine Learning (ICML 2020), 2020.