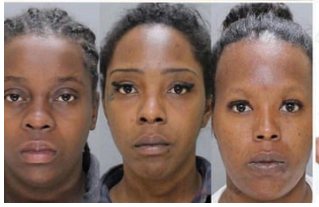
Rubén Martínez Cantín
Dpto. Informática e Ingeniería de Sistemas.

(Decision making and RL Break)

- **Bayesian decision making**
  - Active learning
  - Bayesian optimization

Universidad Zaragoza






vestido de mujer azul y negro

polo de hombre a rayas azules y negras

Universidad Zaragoza

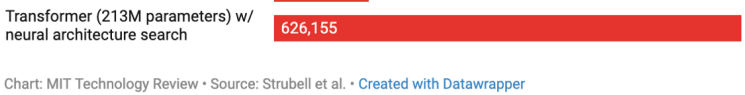
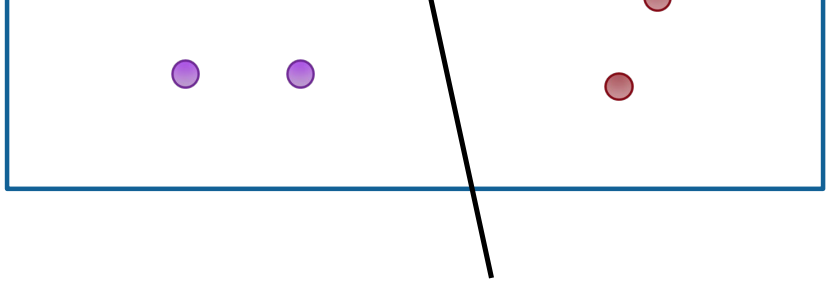
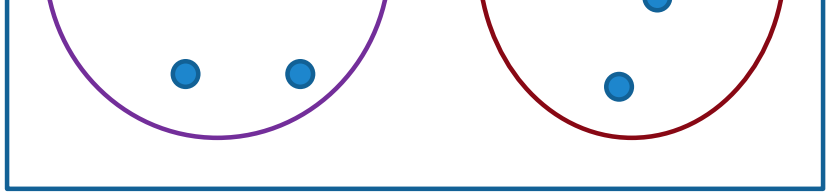


Chart: MIT Technology Review • Source: Strubell et al. • Created with Datawrapper

- We need new paradigm of Small Data.

Universidad Zaragoza

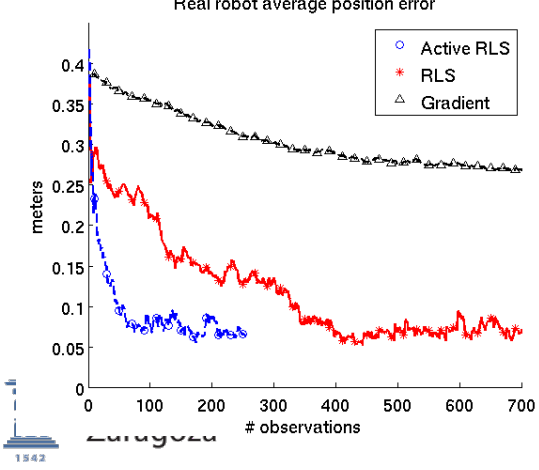


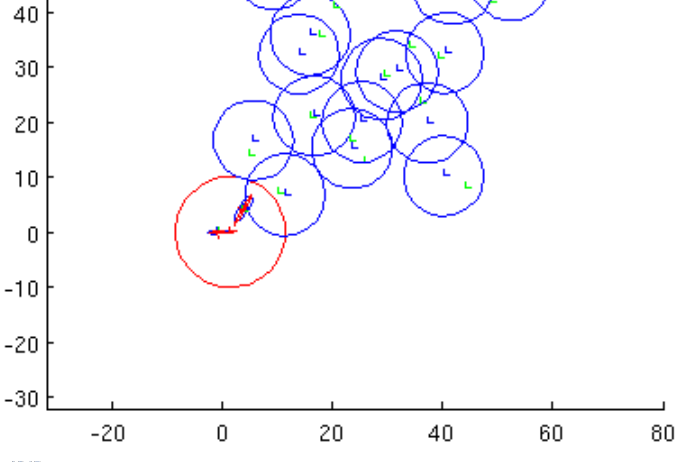


- And we can interact with data

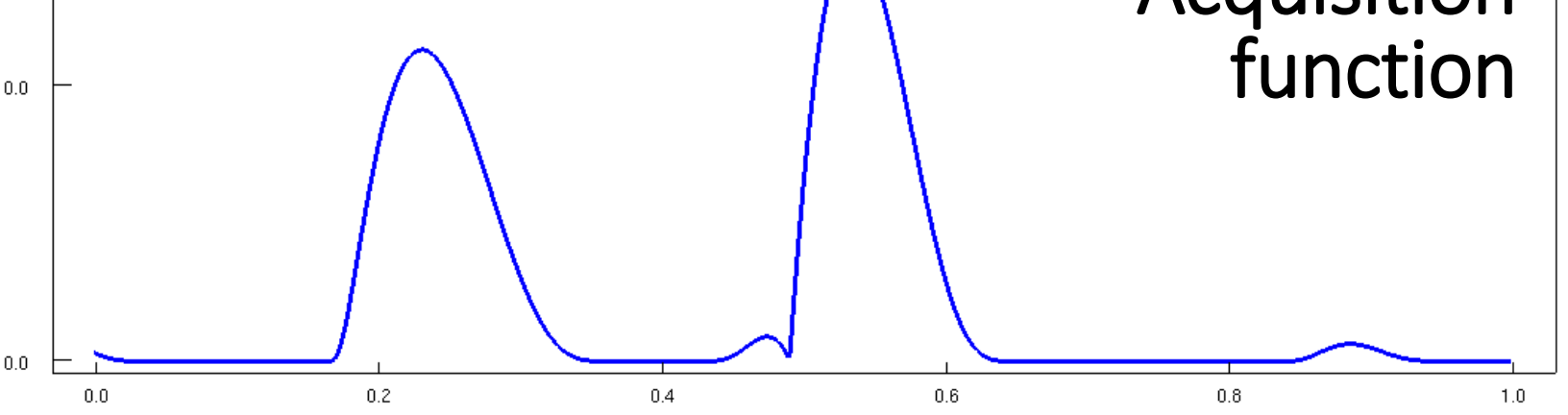
possible data. **Good data is often better than a lot of data.**

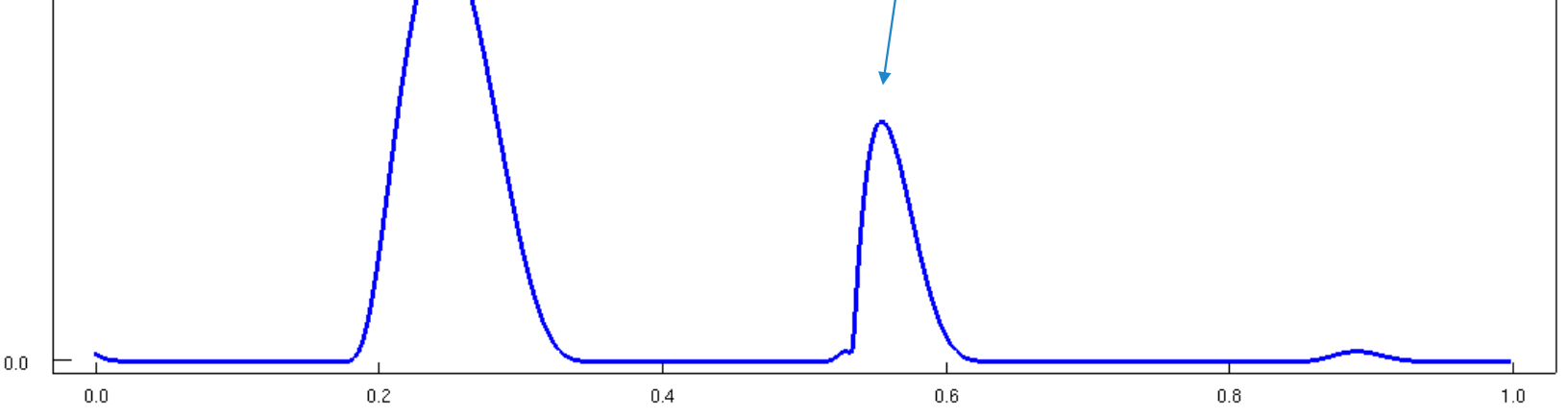
- Humans also query the environment continuously…

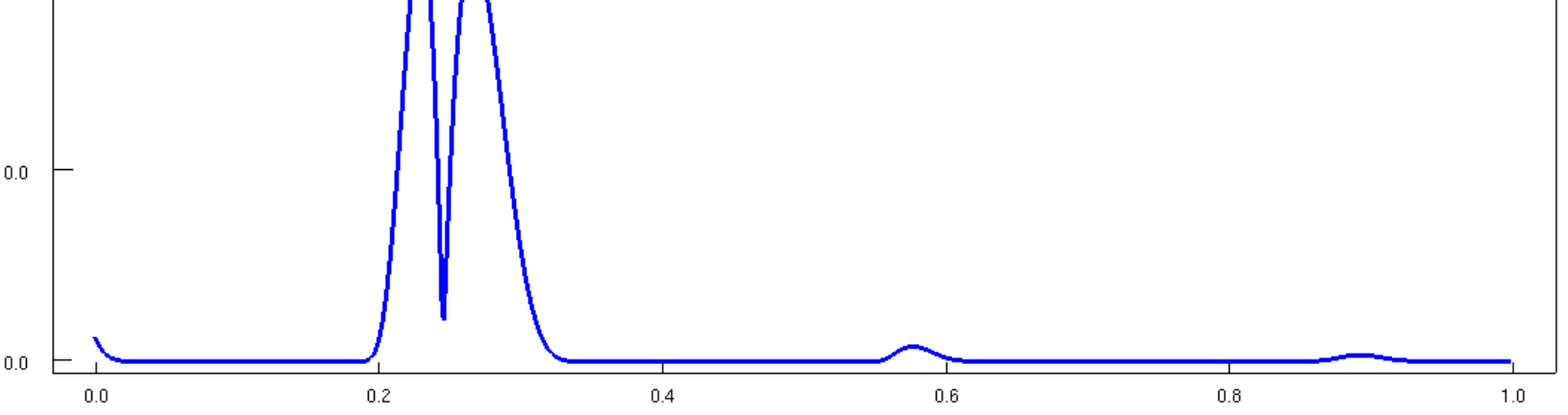

Real robot average position error
Active RLS
RLS
Gradient
meters
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
100
200
300
400
500
600
700
observations


40
30
20
10
0
-10
-20
-30
-20
0
20
40
60
80

Active policy search: Bayesian optimization


Acquisition function
0.0
0.0
0.0
0.2
0.4
0.6
0.8
1.0


0.0
0.0
0.2
0.4
0.6
0.8
1.0


0.0
0.0
0.0
0.2
0.4
0.6
0.8
1.0

- Expected utility → e.g.: Expected Improvement

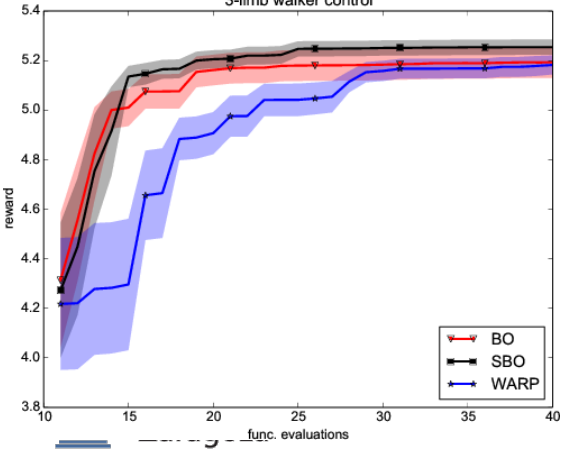
$$\mathbb{E}[I(x)] = \underbrace{(\mu(x) - f(x_{best}))}_{exploit}\,\Phi(d) + \underbrace{\sigma^2(x)}_{explore}\,\phi(d)$$

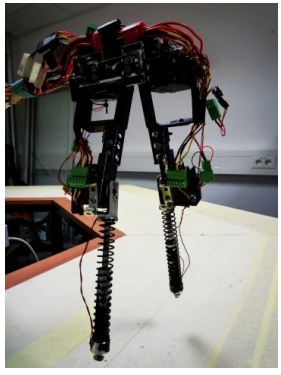
$$d = \frac{x-\mu(x)}{\sigma(x)}$$

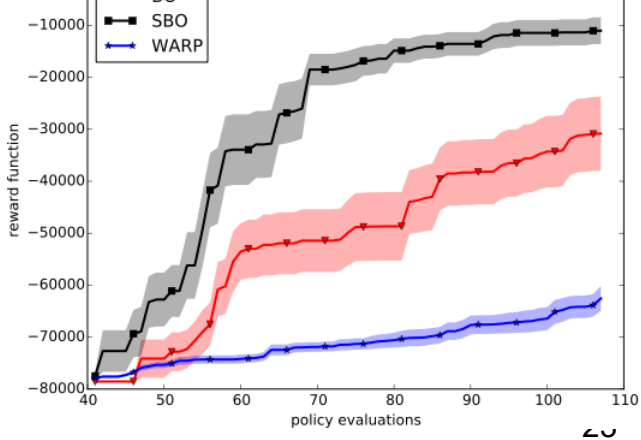
Universidad Zaragoza

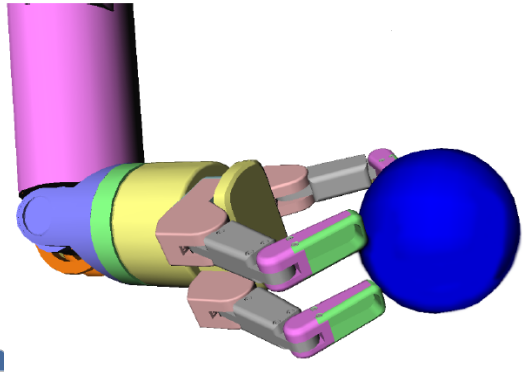
- Augment the data with the new point and response $y_i = f(x_i) + \epsilon$:
  $\mathbf{X} \leftarrow \{\mathbf{X} \cup \mathbf{x}_i\}$ $\mathbf{y} \leftarrow \{\mathbf{y} \cup y_i\}$ $i \leftarrow i + 1$

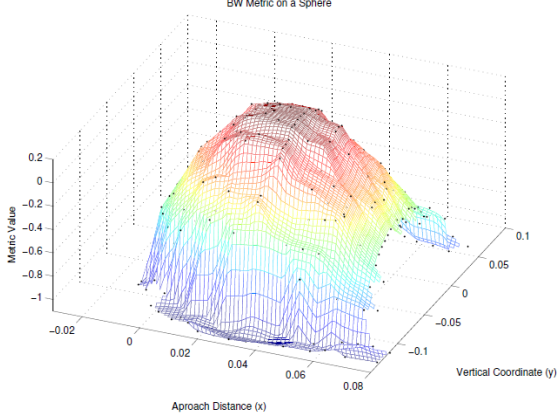

FOR NOTICING THIS
NEW NOTICE
YOUR NOTICING IT
HAS BEEN NOTED
AND WILL BE REPORTED TO THE AUTHORITIES

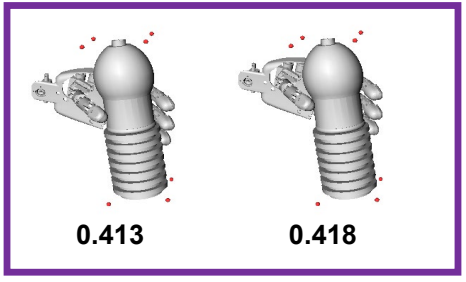

3-limb walker control
reward
func. evaluations
BO
SBO
WARP

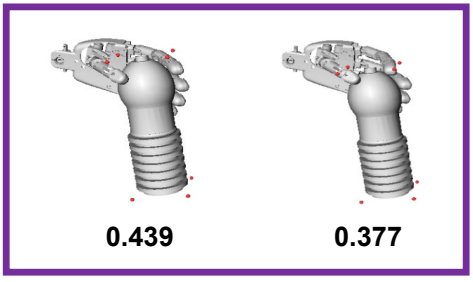



reward function
policy evaluations
BO
SBO
WARP



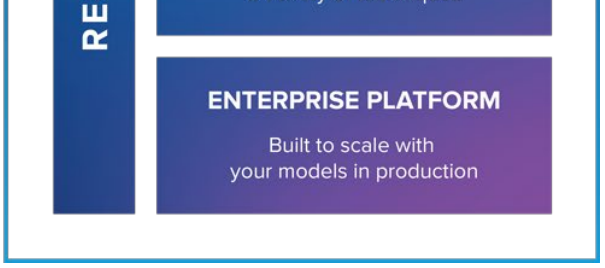

BW Metric on a Sphere
Metric Value
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
−0.02
0
0.02
0.04
0.06
0.08
Aproach Distance (x)
0.1
0.05
0
−0.05
−0.1
Vertical Coordinate (y)

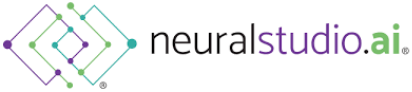


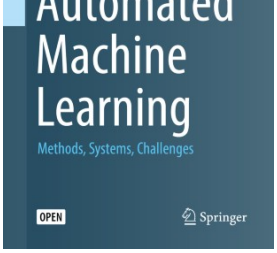


Universidad Zaragoza
1542


Testing
Data
Model
Evaluation
Objective Metric


ENTERPRISE PLATFORM
Built to scale with
your models in production


neuralstudio.ai


Automated
Machine
Learning
Methods, Systems, Challenges
OPEN
Springer


AutoWeka
An Automated Data Mining Software Based on Weka


TPOT


bigml


Xpanse
AI


OPTUNA


AutoKeras


tune


Universidad
Zaragoza
1542


mind
FOUNDRY

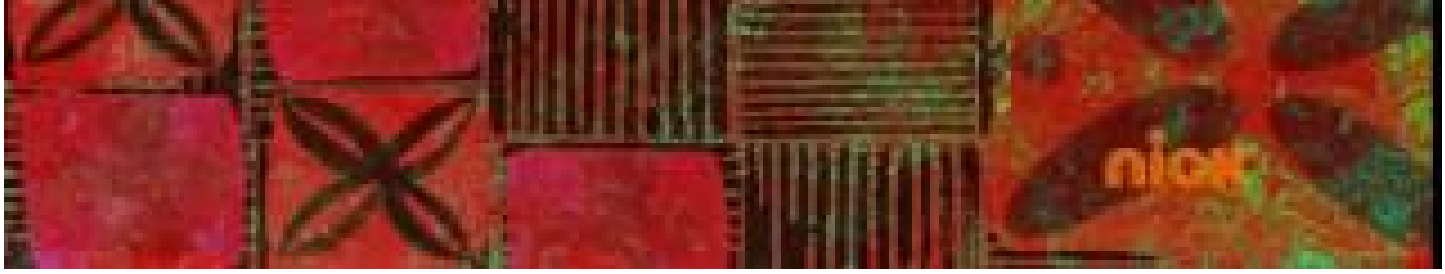
Bayesian Optimization in AlphaGo

Yutian Chen, Aja Huang, Ziyu Wang, Ioannis Antonoglou, Julian Schrittwieser, David Silver & Nando de Freitas

DeepMind, London, UK
yutianc@google.com

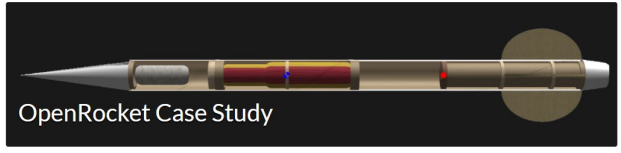
Abstract

During the development of AlphaGo, its many hyper-parameters were tuned with Bayesian optimization multiple times. This automatic tuning process resulted in substantial improvements in playing strength. For example, prior to the match with Lee Sedol, we tuned the latest AlphaGo agent and this improved its win-rate from 50% to 66.5% in self-play games. This tuned version was deployed in the final match. Of course, since we tuned AlphaGo many times during its development cycle, the compounded contribution was even higher than this percentage. It is our hope that this brief case study will be of interest to Go fans, and also provide Bayesian optimization practitioners with some insights and inspiration.

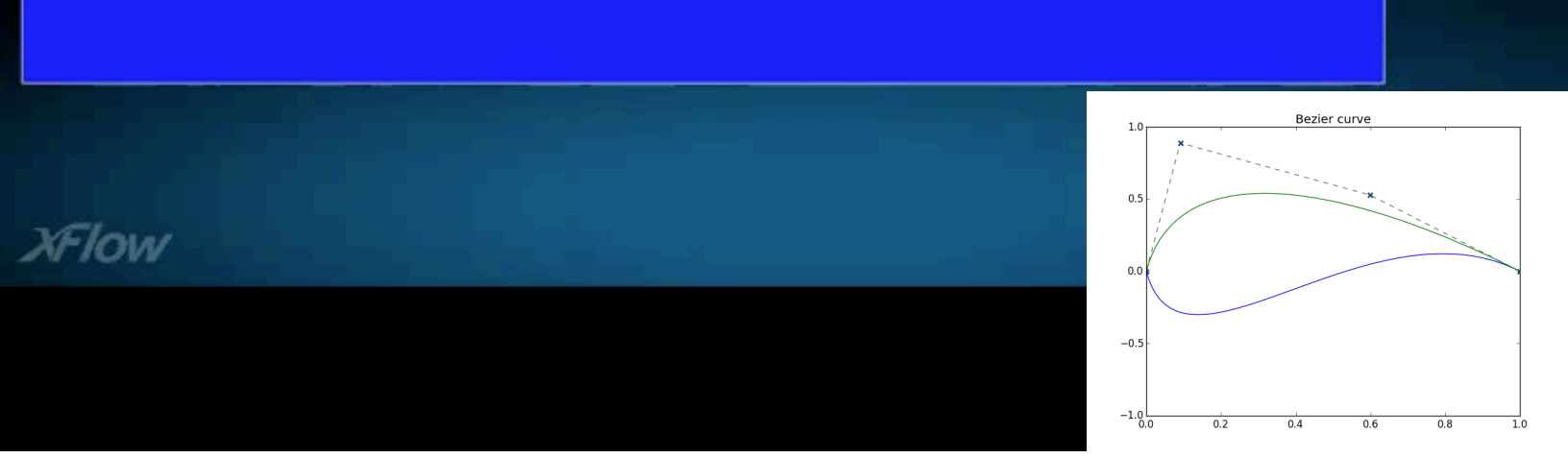


SIGOPT

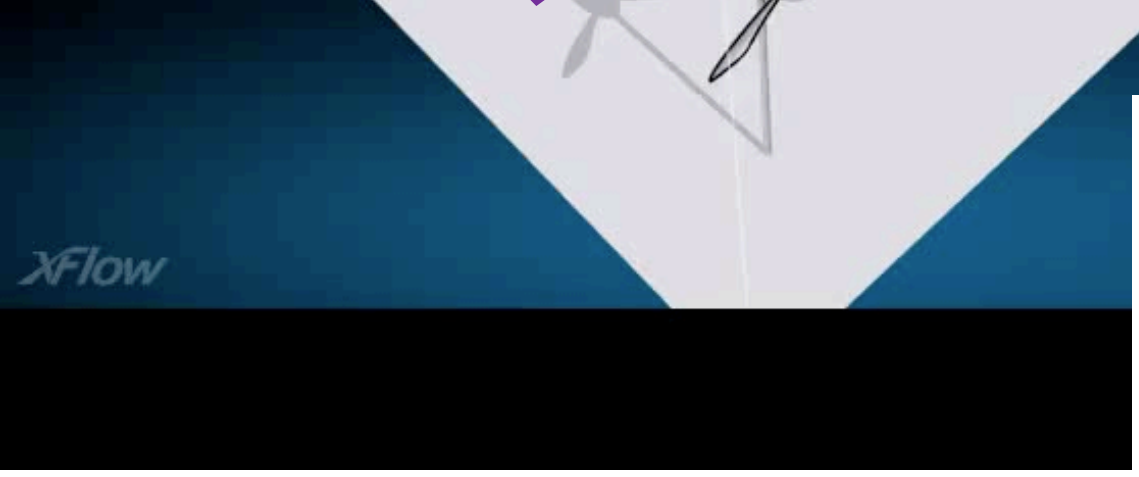
NASA

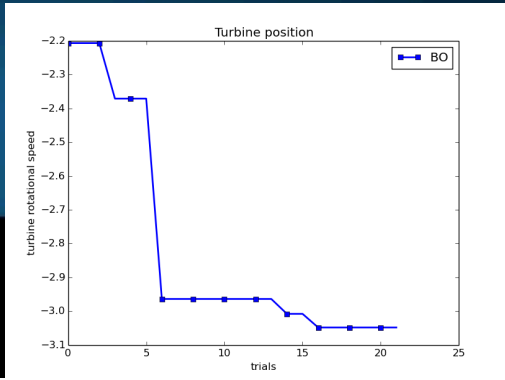


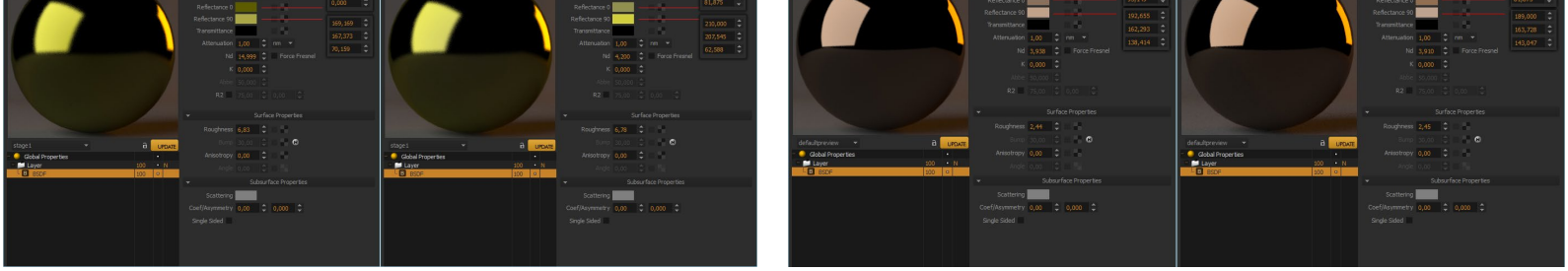
Universidad Zaragoza

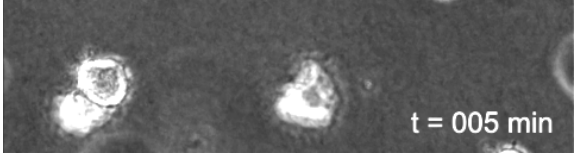

XFlow

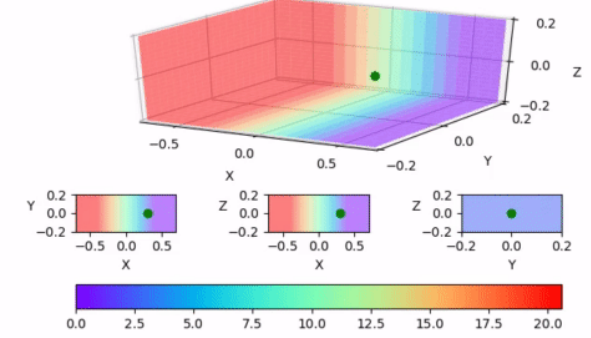
Bezier curve
1.0
0.5
0.0
−0.5
−1.0
0.0
0.2
0.4
0.6
0.8
1.0


XFlow

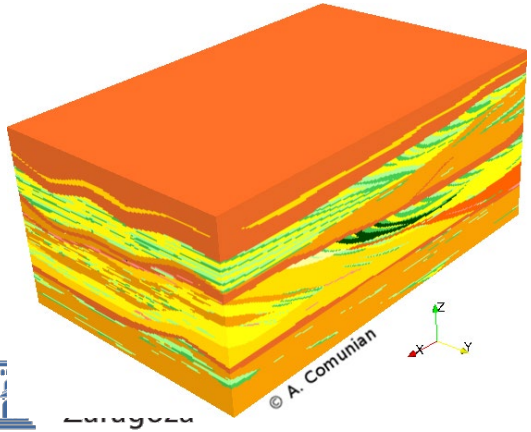

Turbine position
BO
turbine rotational speed
trials
−2.2
−2.3
−2.4
−2.5
−2.6
−2.7
−2.8
−2.9
−3.0
−3.1
0
5
10
15
20
25



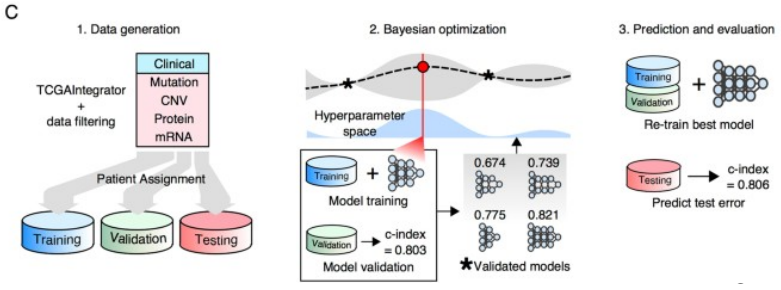




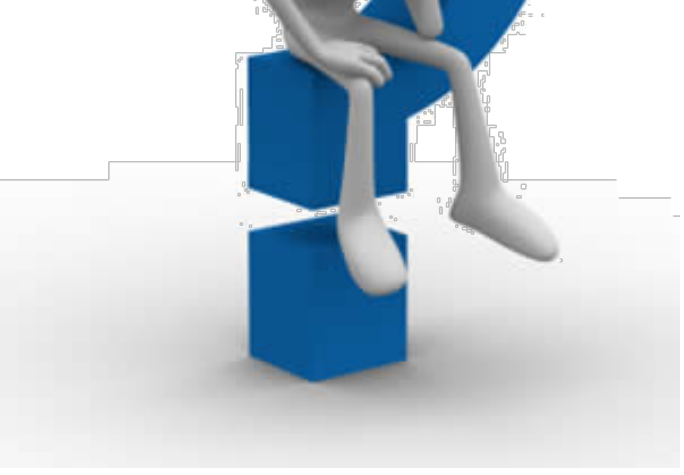
Universidad Zaragoza


Z
X
Y
© A. Comunian


Molecular features
diagnosis (time = 0)
time
better?
Prediction error (negative log-likelihood)


C
1. Data generation
TCGAIntegrator + data filtering
Clinical
Mutation
CNV
Protein
mRNA
Patient Assignment
Training
Validation
Testing
2. Bayesian optimization
Hyperparameter space
Model training
Model validation
c-index = 0.803
0.674
0.739
0.775
0.821
Validated models
3. Prediction and evaluation
Re-train best model
Predict test error
c-index = 0.806

Zaragoza



and Computer Vision