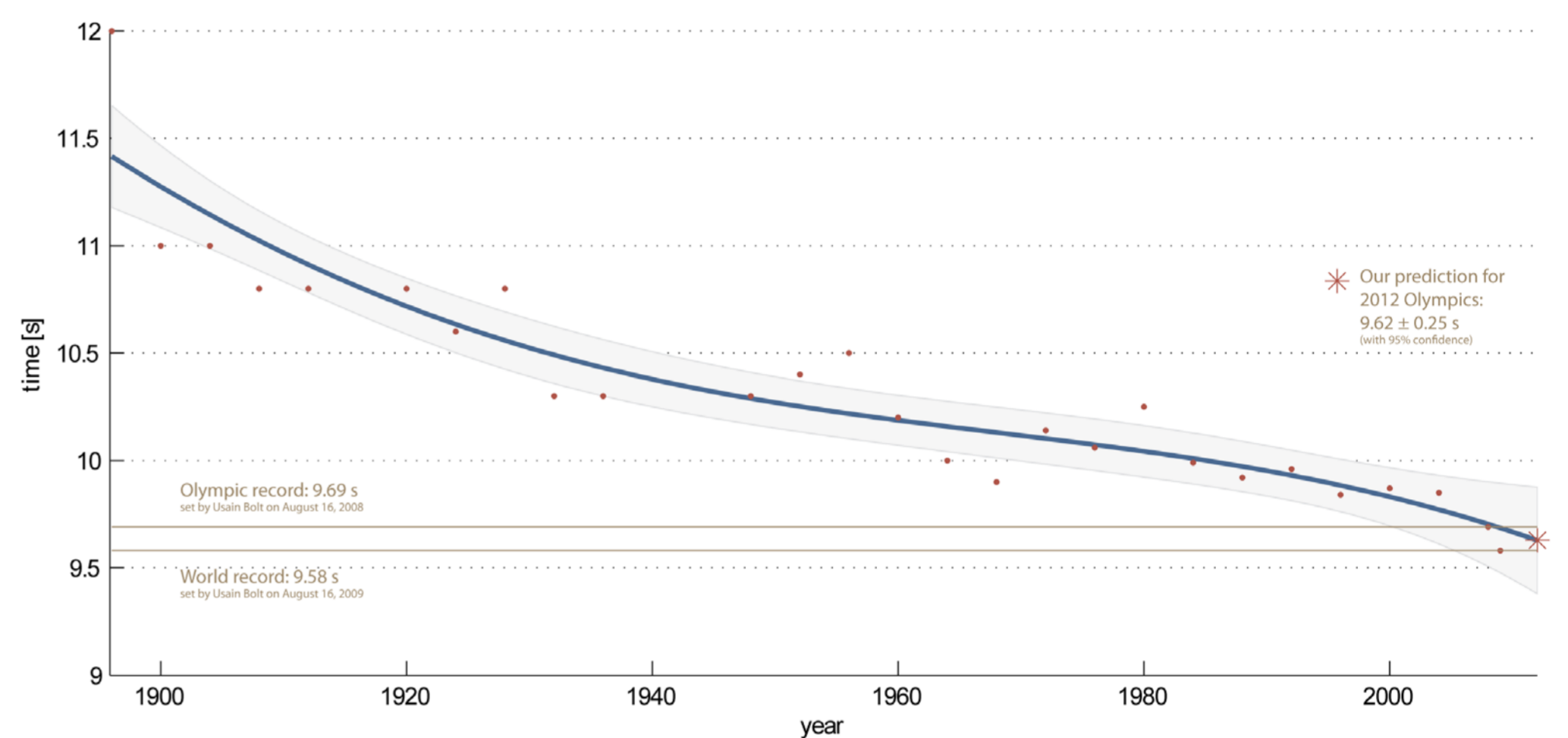


GAUSSIAN PROCESSES FOR MACHINE LEARNING

Dejan Petelin, univ. dipl. inž. rač. in inf.
Study programme: Information and Communication Technologies,
Jožef Stefan international postgraduate school
Mentor: prof. dr. Juš Kocijan
Institut Jožef Stefan, Jamova cesta 39, 1000 Ljubljana

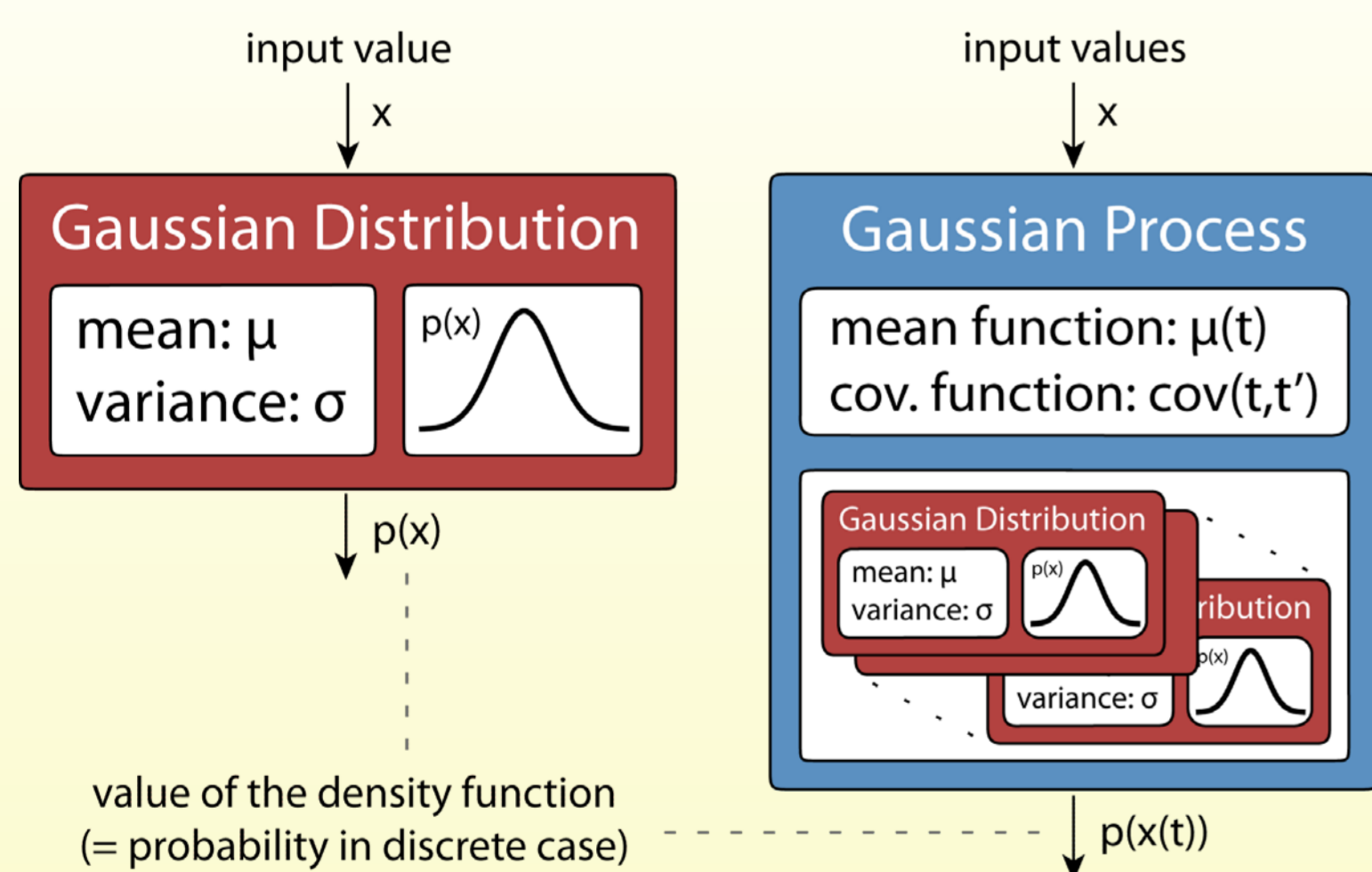


**Will the 100m
sprint record
be broken at the
2012 Olympics?**



BASIC INFORMATION

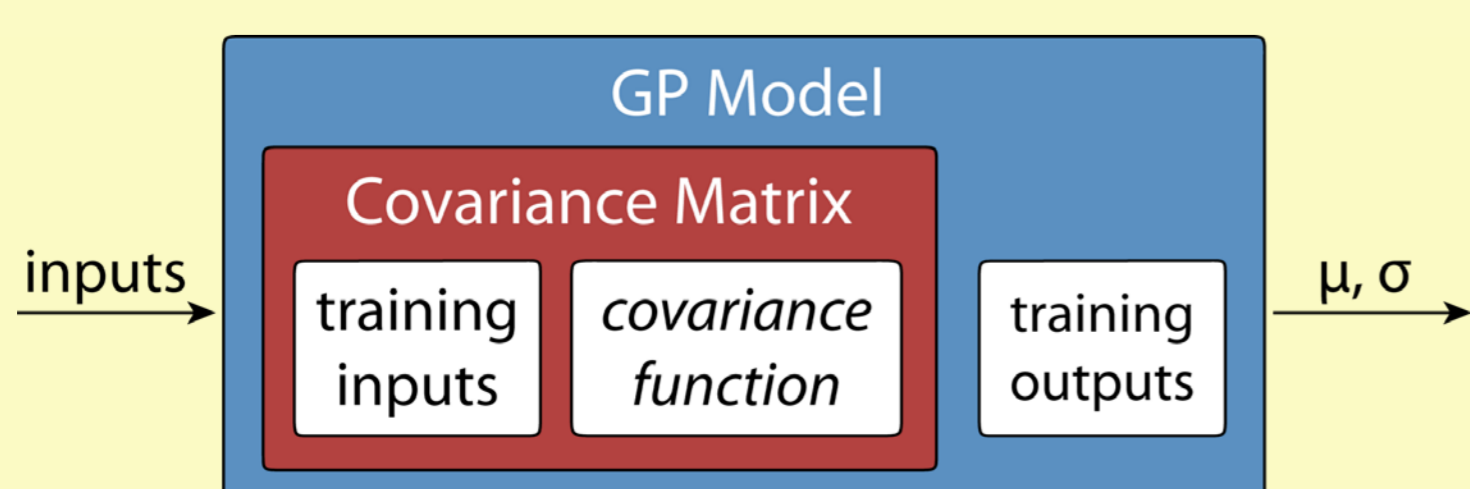
- GP model is non-parametric,
- probabilistic (Bayesian) method,
- it may be considered as a Black-box method while
- it searches relationships among data.
- It is a kernel method.



- It is easy to use,
- returns also a confidence in a prediction,
- is resistant to missing and noisy data and
- can't be overfitted (due to Bayesian nature).
- It has ability to include prior knowledge.

MODELING

- GP model is fully specified with the data and a covariance matrix

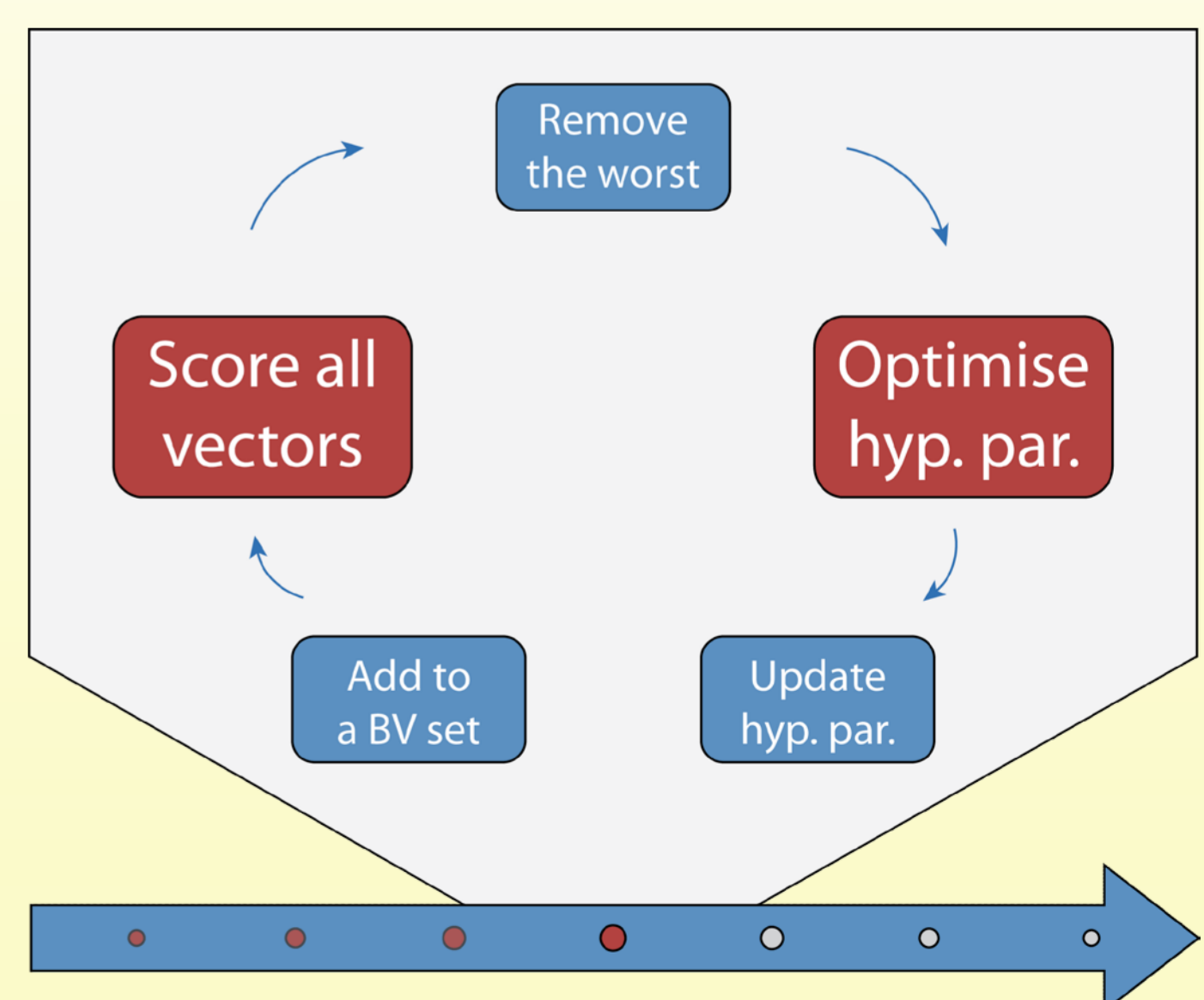


COVARIANCE FUNCTION

- It reflects correlation between inputs,
- must generate semi positive-definite matrix.
- It is constructed out of a number of *hyperparameters*
- that may be used to adjust the properties of the GP prior

ONLINE LEARNING

- Improving a model with every new data.
- Maintaining a subset of the most informative data (BV).
- Enables forgetting of old (irrelevant) data



REAL APPLICATIONS



Wastewater systems



Weather (ozone, CO₂,...)



Robotics

- and many others