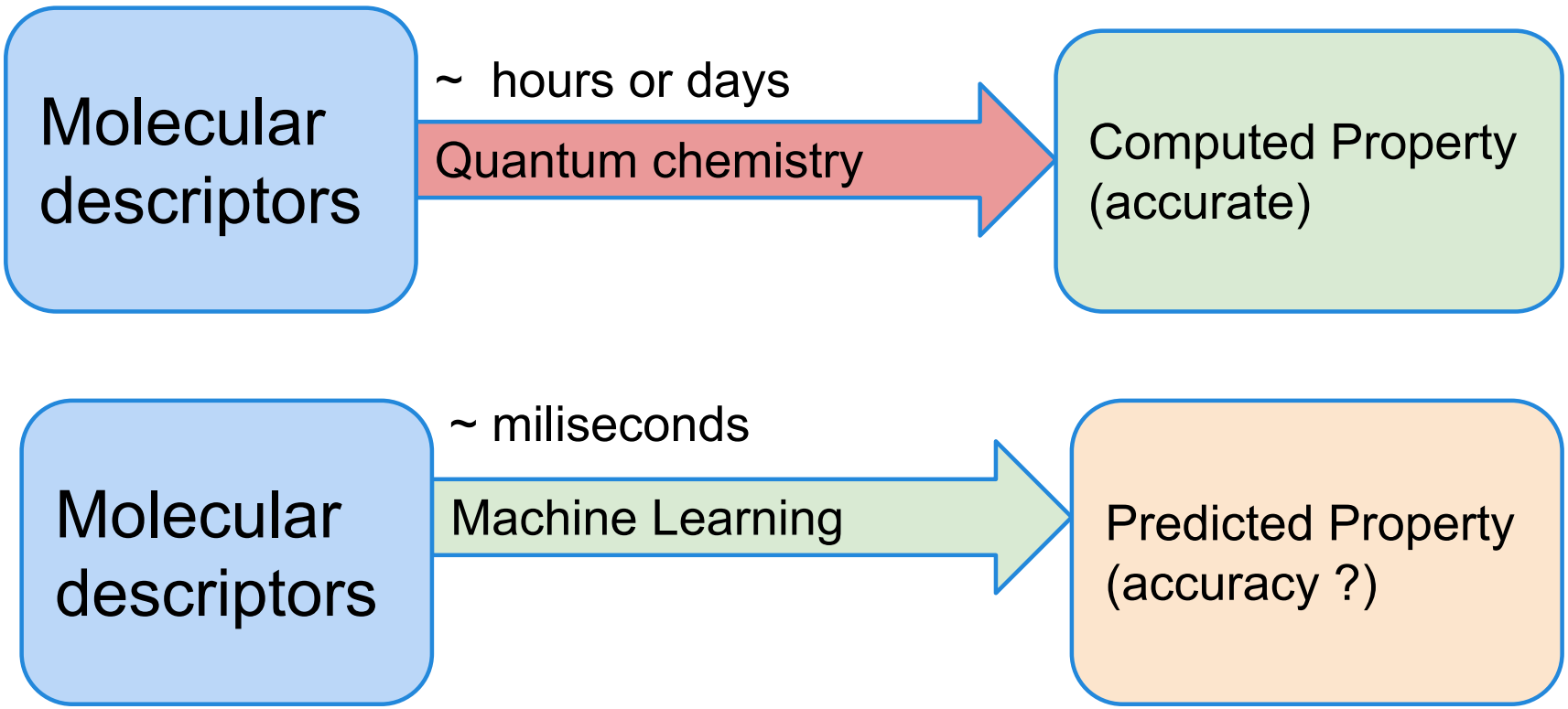
# Machine Learning techniques for predicting molecular properties

Viviana Petrescu

July 1, 2015

# Overview

- Compute/predict properties of molecules for materials design
  Examples: drug discovery, water purification, energy transmission and storage
- Quantum Chemistry calculations are expensive
- Machine Learning could predict properties of molecules at a fraction of the cost


Molecular descriptors
~ hours or days
Quantum chemistry
Computed Property (accurate)
Molecular descriptors
~ miliseconds
Machine Learning
Predicted Property (accuracy ?)

# Overview

# Molecular descriptors

- Desired properties of Molecular Descriptor
  - invariance to atom indexing
  - invariance to rotation and translation
- Coloumb Matrix [*Rupp*2012] scales $O(N^2)$

$$C(i,j) = \begin{cases} 0.5 * Z_i^{2.4} & i == j \\ Z_i * Z_j / ||R_i - R_j|| & i \neq j \end{cases} \quad (1)$$

$Z_i$ is charge of atom i
$R_i$ is 3D position of atom $i$

# Coloumb Matrix

Coloumb Matrix descriptor

- invariant to rotation and translation (use $||R_i - R_j||$ )
- invariance to atom indexing
  - Sorted Coloumb Matrix - indexes given by sorting the row norms
  - Random Coloumb Matrix - generate multiple Sorted Coloumb Matrices perturbed by noise)

# Random Coloumb Matrix

Coloumb Matrix descriptor

- Atomization energy prediction for a dataset of 7k samples (5.5k training 1.5k testing) with H,O,C,N,S
  Multilayer Feed Forward NN with Random Coloumb
  3.1 kcal/mol MAE - chemical accuracy level 1kcal/mol

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Introduction

Things in a Bulleted List

- Bullets that
- Come up
- One by one

# Overview