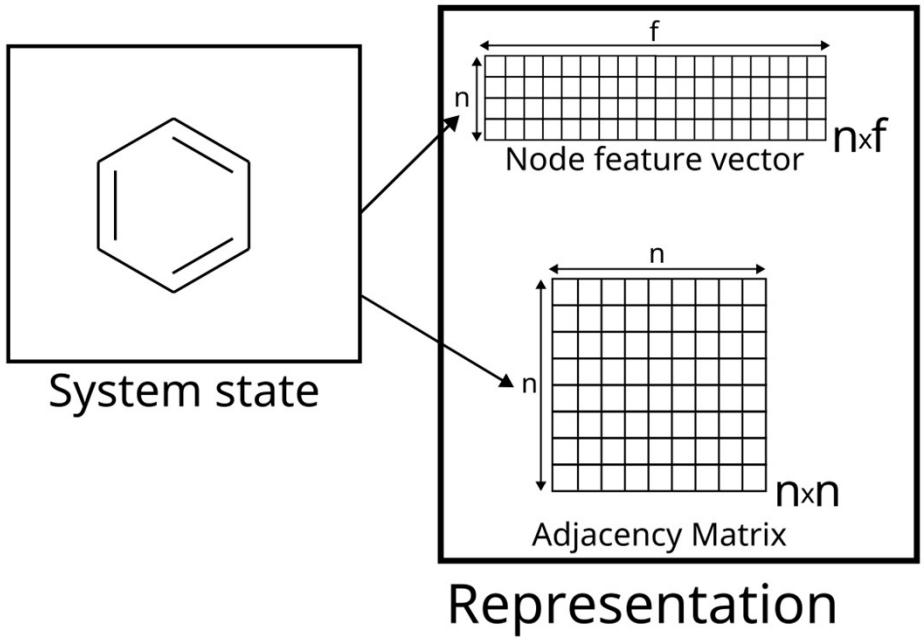
# Deep Assortment of Molecular Property (DAMP)

**Venkata Sai Sreyas Adury and Arnab Mukherjee**

# Molecules as graphs

**All molecules can be represented as graphs**

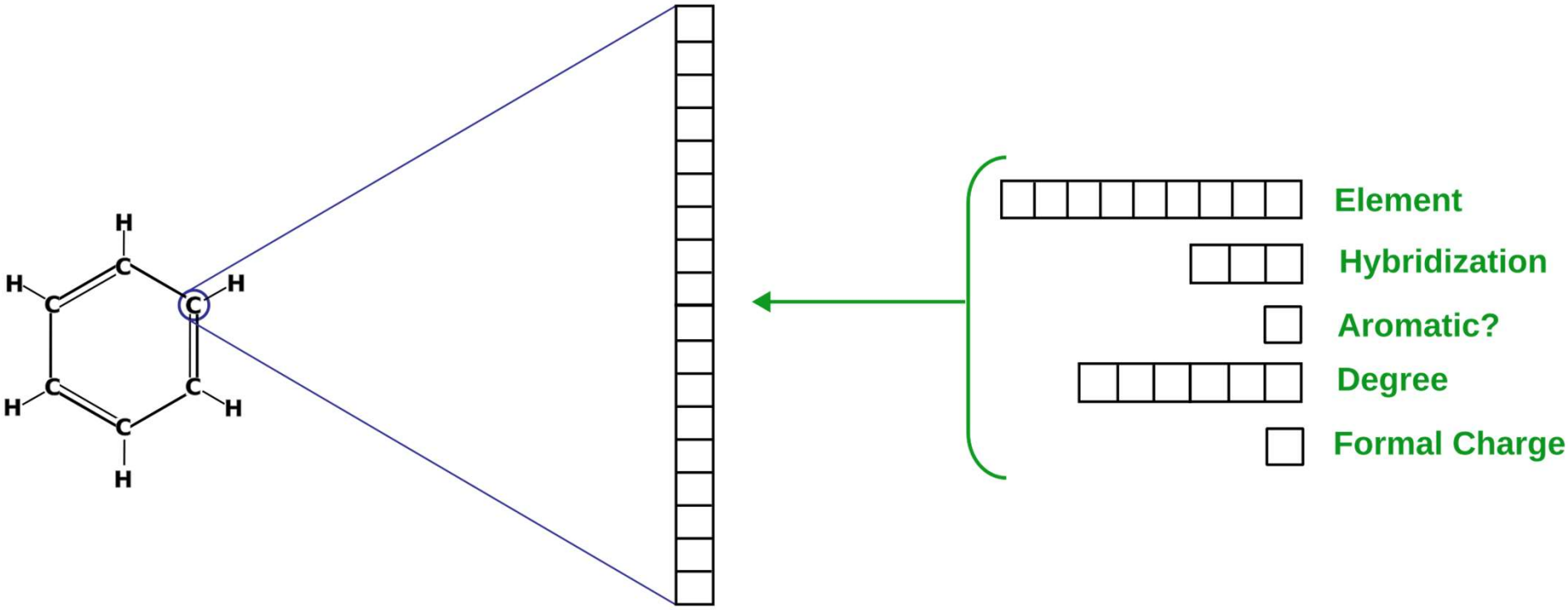
- The atoms form the nodes
- The bonds form the edges




The node *features* can store information about a node (i.e. atom)

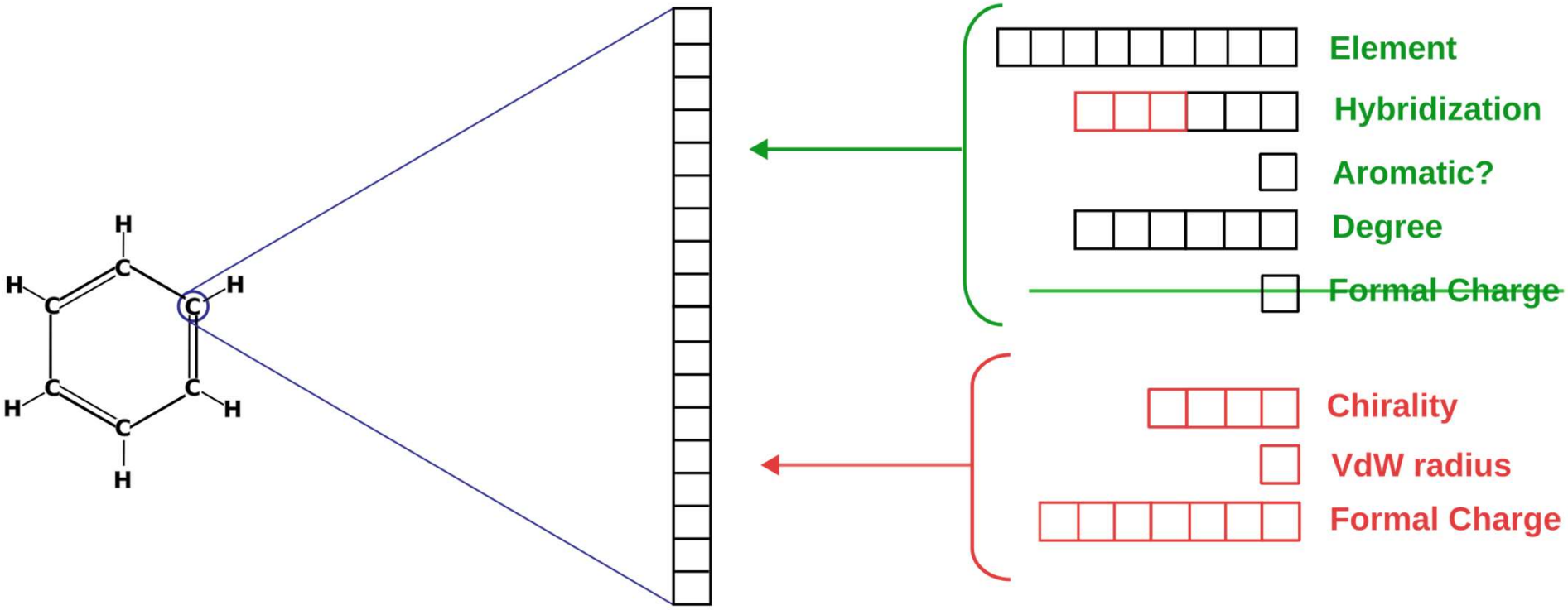
# Node featurization

**We use DeepChem's featurization (feature-size=30)**



# More features

**For some properties, we added more features (final feature size = 79)**



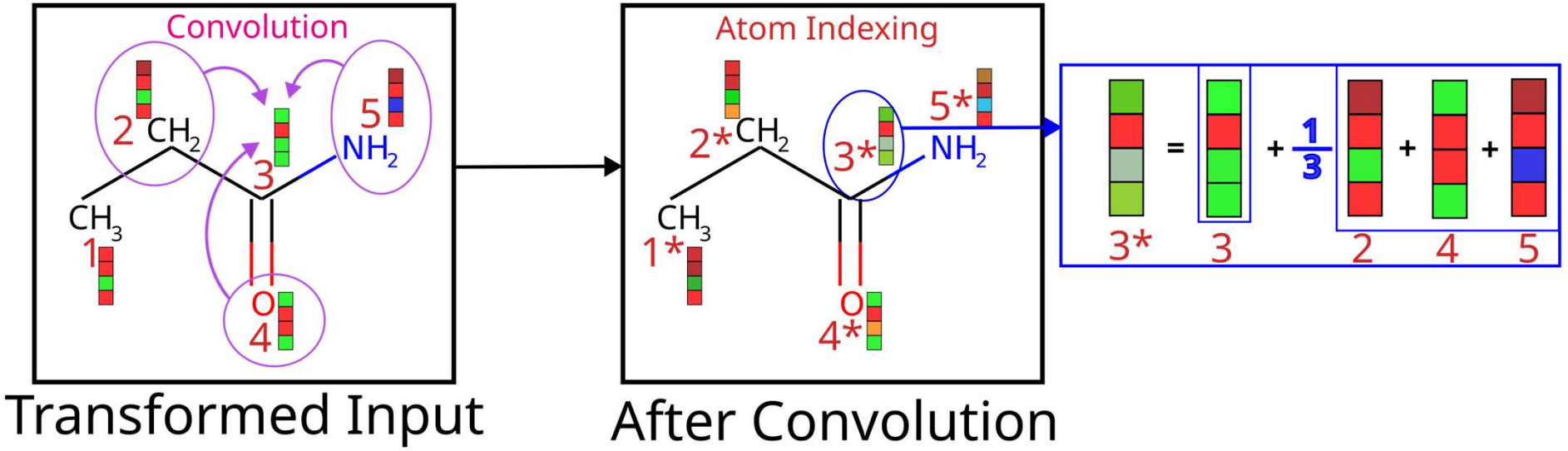
# Changes to the features

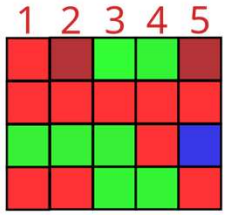
- Expanded the number of elements possible
- Used one-hot encoding (-3, -2, -1, 0, 1, 2, 3) for formal charge instead of a single integer
- Added potentially important chemical parameters such as Vdw radius and covalent radius

# Graph Convolution and our model

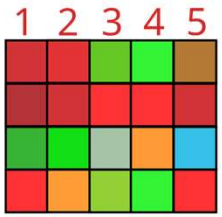
- Convolution lets us enrich the information available at each node by incorporating it from its nearest neighbours
- Multiple iterations of GCN can thus incorporate information from farther and farther neighbours
- In out models, we used 8 iterations of Graph Convolution and trained with the Adam optimizer after taking an 80-20 split of the provided data for training and validation respectively

# Graph Convolution



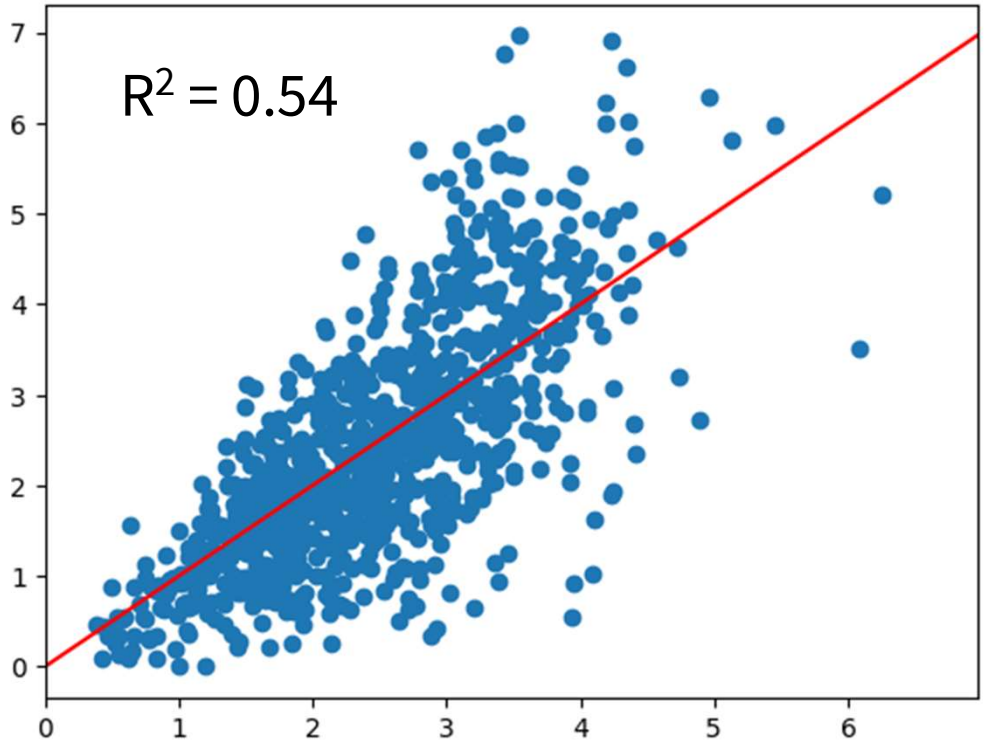




Initial features

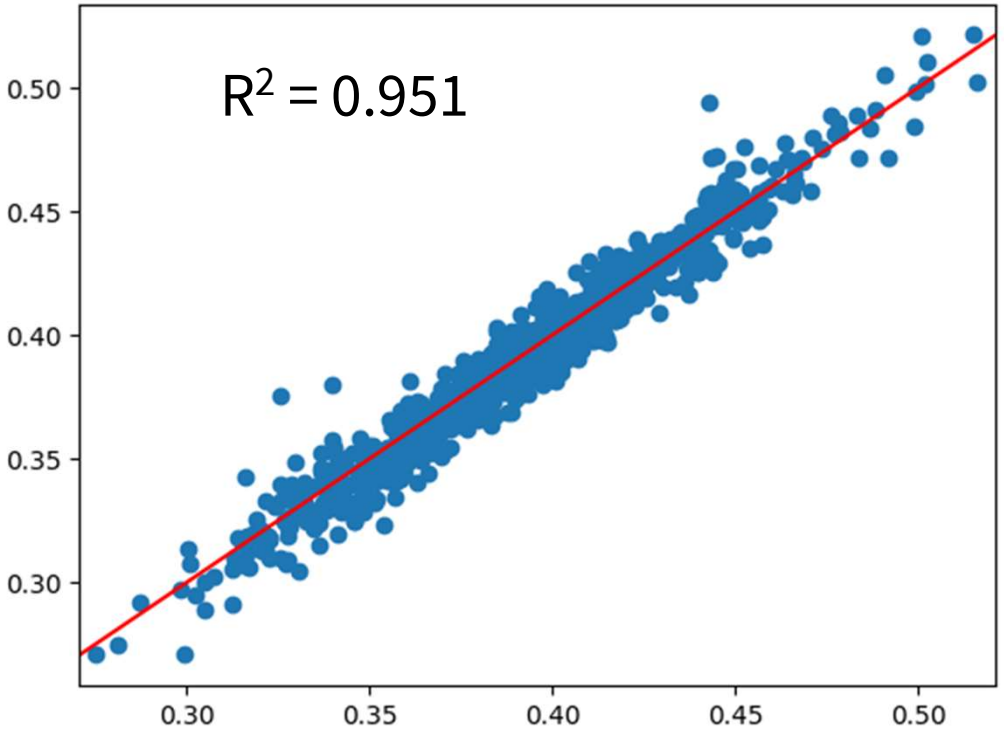



Final features

# Results (Validation set)

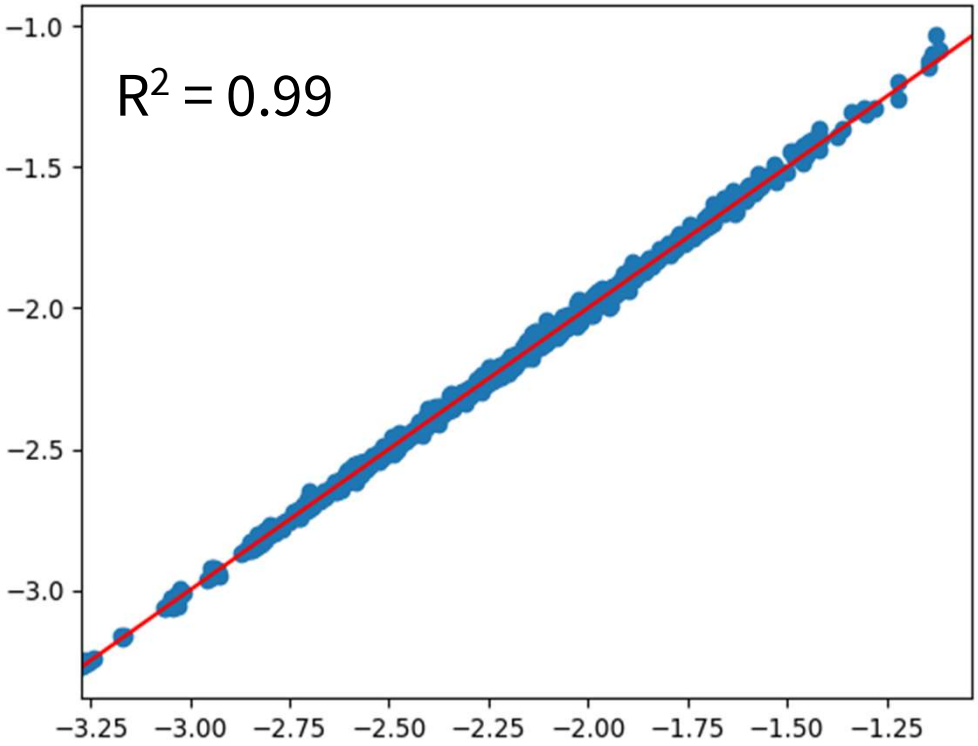



Property 1

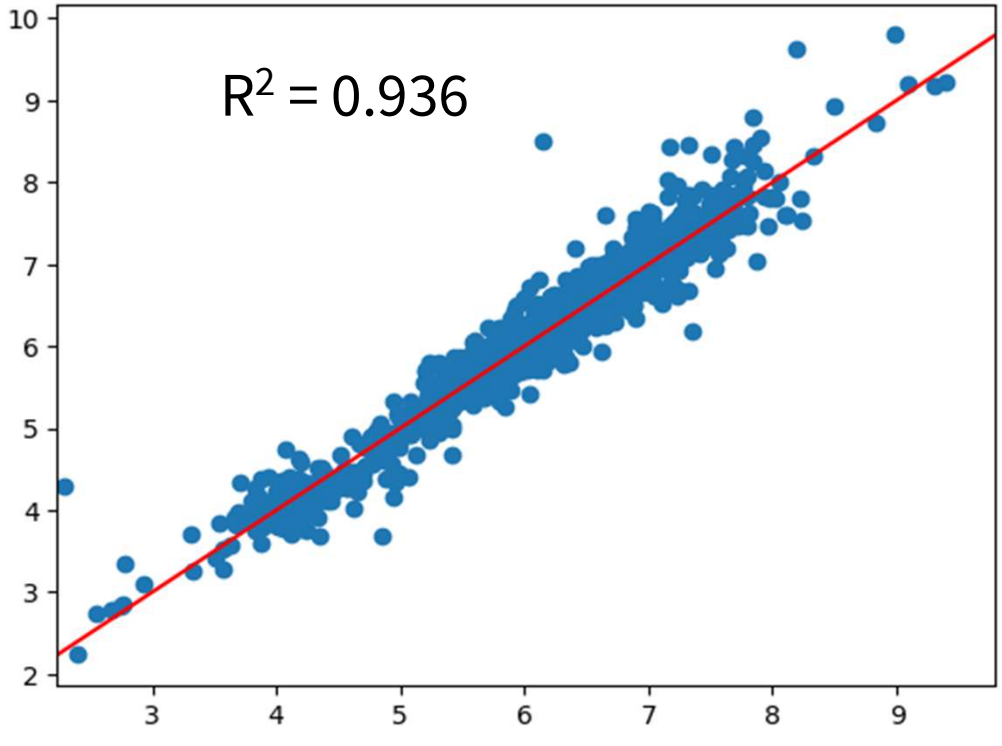



Property 2

# Results (Validation set)




Property 3




Property 4