- Sources
  - CS229 notes (8-15)
  - https://www.analyticsvidhya.com/blog/2021/05/all-you-need-to-know-about-your-first-machine-learning-model-linear-regression/
  - https://medium.com/@satyavishnumolakala/linear-regression-pros-cons-62085314aef0
- 2 types - simple and multiple
- y is a linear function of the parameters
- while gradient descent in general is susceptible to local minima, the optimization problem posed for linear regression has only one global optima, thus, gradient descent always converges (assuming the learning rate is not too large) to the global minimum
- **Input**
  - vector of features
- **Output**
  - continuous (for all regression problems)
- **Advantages**
  - Simple
  - Computationally efficient
  - Interpretability of the output
- **Disadvantages (can be avoided by using regularization, data preprocessing, and dimensionality reduction)**
  - Too simplistic
  - Severely affected by outliers
  - Many assumptions
- **Evaluation metrics**
  - R-squared
  - Adjusted R-squared
  - Mean squared error (MSE)
  - Root mean squared error (RMSE)