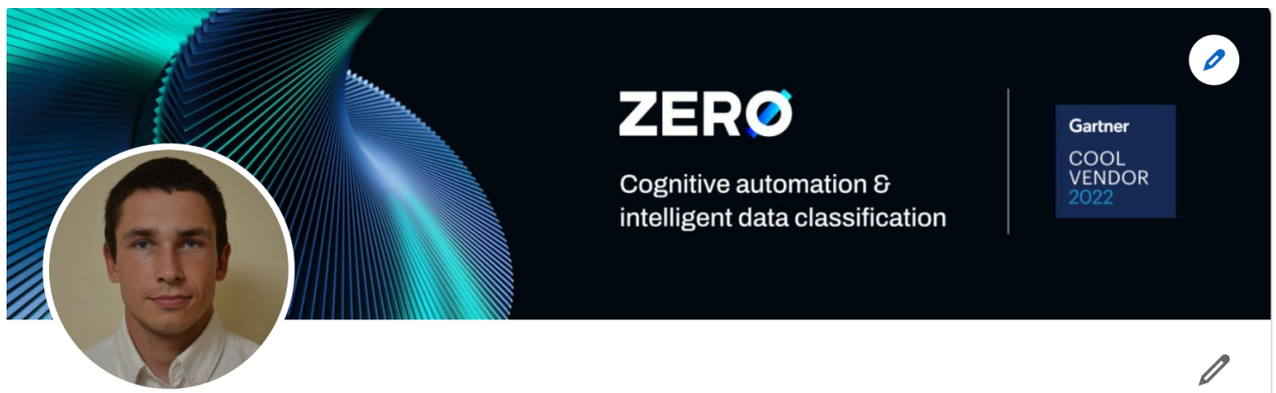
Deep Learning for NLP

Sorokin Semen

Based on MS DL Course of Boris Zubarev @bobazooba

Introduction



Semyon Sorokin
TL Data Science
Yerevan, Armenia · Contact info

ZERO Armenia
Higher School of Economics

Experience

Lead NLP Engineer
ZERO Armenia · Full-time
Nov 2022 - Present · 11 mos
Armenia · On-site

Leading team of web Python developers, data scientists, and annotation specialists. Creating a product for document processing. Automate document classification and entity extraction. ML engine is lear ...see more

Skills: Deep Learning · Natural Language Processing (NLP) · Named Entity Recognition · Document Classification

Lecturer
Higher School of Economics · Part-time
Mar 2019 - Present · 4 yrs 7 mos
Moscow, Moscow City, Russia

1. Conducting lectures and seminars of the course "Neural networks in NLP tasks" using PyTorch
- Text classification task (based on LSTM, CNN; self-attention)... ...see more

Skills: Machine Learning · Text Classification · Deep Learning · Python (Programming Language)

Data Scientist
Sberbank · Full-time
Feb 2020 - Nov 2022 · 2 yrs 10 mos
Moscow, Moscow City, Russia

NLP Team Lead. A full cycle of project management (from DS point of view) for building services using deep learning. Working with text recognition models for a document photos/scans. Tasks: extraction o ...see more

Skills: Machine Learning · Text Classification · Deep Learning · Python (Programming Language) · Natural Language Processing (NLP) · Transformer · BERT · Named Entity Recognition · Document Classif ...see more

Grading system

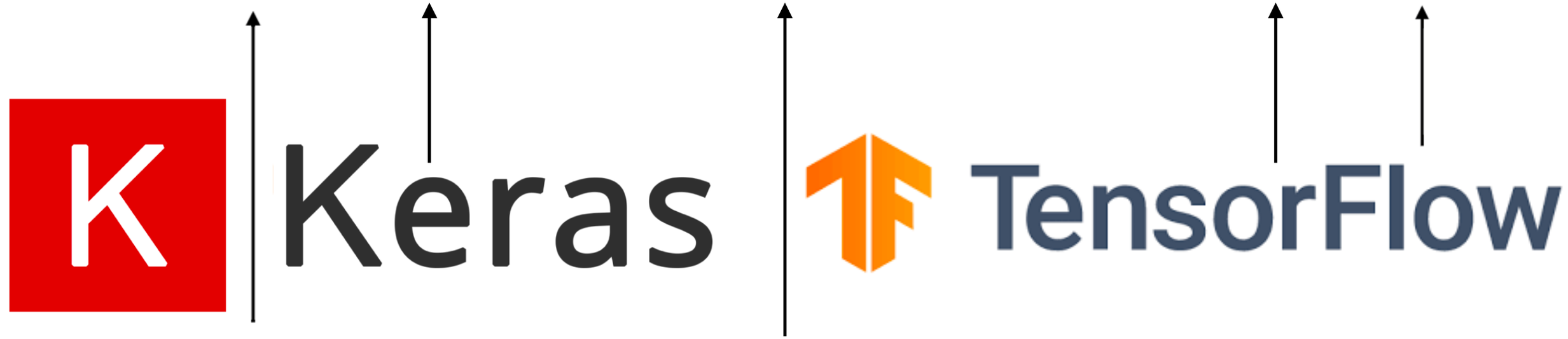
Three labs	**40% of cumulative assessment**
Three / five (not decided yet) test	**30% of cumulative assessment**
Final project	**30% of cumulative assessment**

If cumulative assessment >= 6:
free to go or increase your score on the exam
else:
pass the exam (a result of the exam >= cumulative assessment)

Exam - two topics from different lectures or
(Mb, if you are lucky) questions like "What is the name of the lecturer?" or "Classify (manually) this image"

Framework

Simple **Hard**



OpenAI
GPT - 4

PyTorch

Deep Learning

Classic Machine Learning

Deep Learning

Prompt Engineering

Stellantis South America

Posted on: **13 September, 2023**

- Full Time

Apply

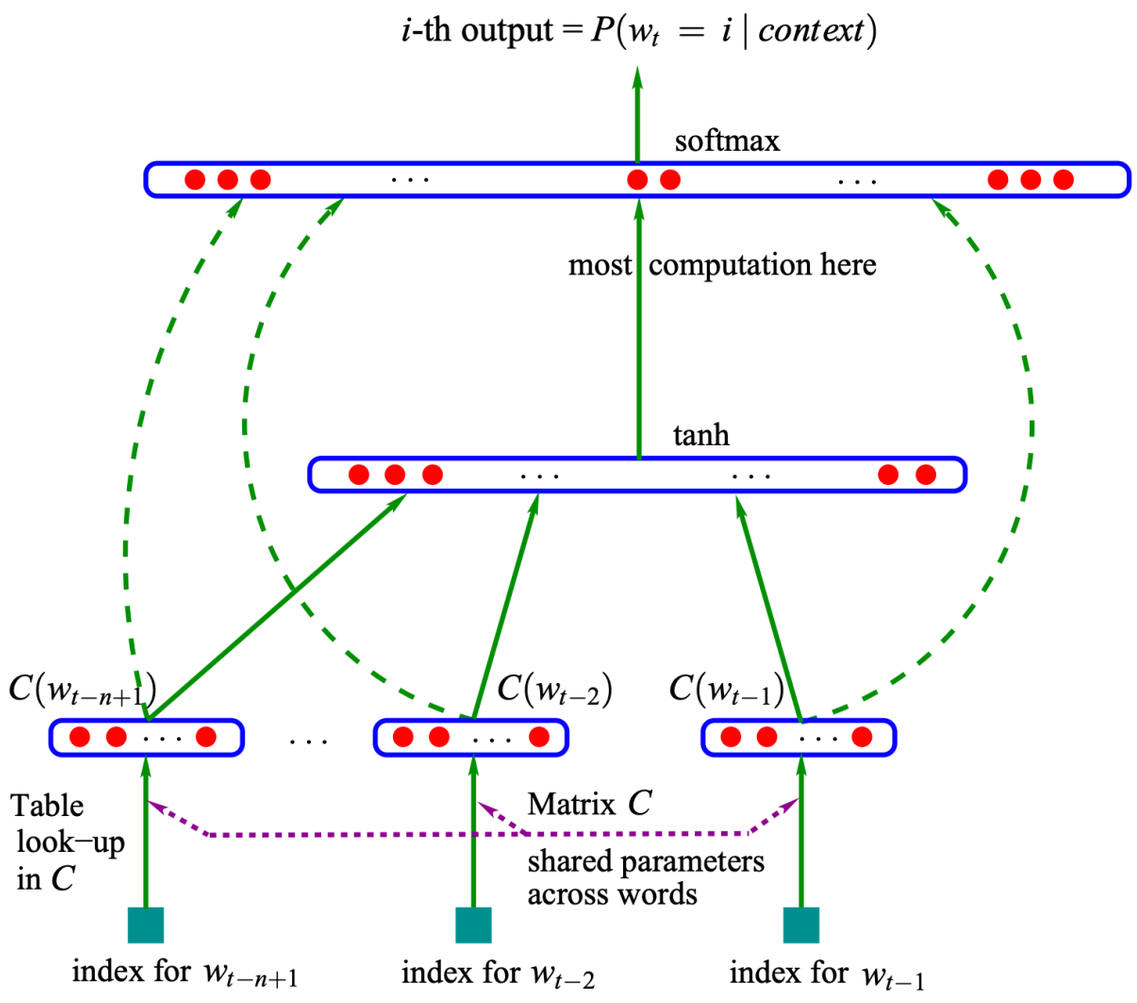
Prompt Engineer

As Prompt Engineer / Generative AI Engineer, your role is to design, develop, refining and optimize AI-generated text prompts to ensure they are accurate, engaging, and relevant for various applications.
It includes natural language processing (NLP) models and prompts that drive the performance and effectiveness of language models and conversational AI systems.
Generative AI Engineer, you will work with generative models and doing prompt engineering to create new and innovative AI products.
You should possess experience in working with Large Language Models (LLMs), utilizing external models, and have the ability to fine-tune open-source models according to the specific requirements of our company.
Collaboration with data scientists, machine learning engineers, and cross-functional teams will be crucial as you focus on creating high-quality prompts, refining model outputs, and enhancing the overall user experience.
Your expertise in NLP algorithms, model engineering, and prompt engineering techniques will play a vital role in shaping the capabilities and performance of AI language models.
PLEASE SHARE YOUR CV IN ENGLISHKey Responsibilities: Prompt Engineering- Design and develop high-quality prompts and templates that guide the behavior and responses of language models. Craft prompts to elicit specific information or control the model's output, ensuring desired accuracy, relevance, and language fluency. Optimize prompts to improve user interactions and system performance. NLP Model Development- Design and develop NLP models, algorithms, and architectures to solve complex language understanding and generation problems.

Deep Learning

Neural Probabilistic Language Model

2003



BERT

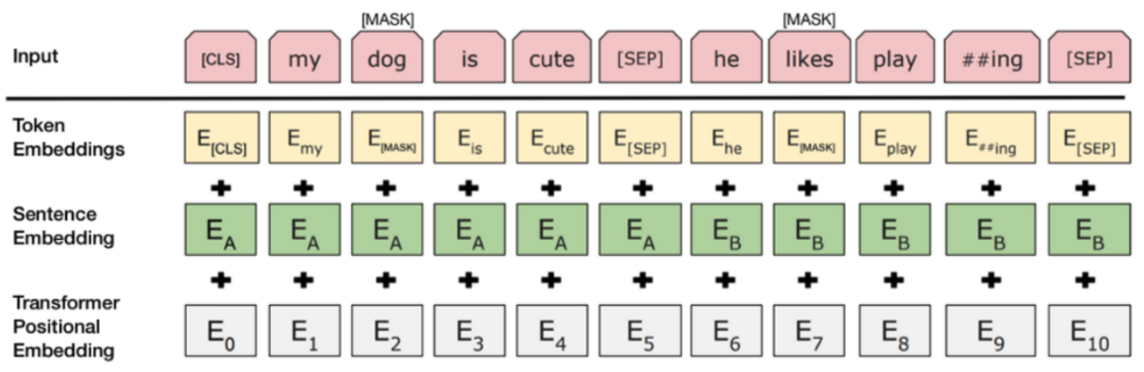
Q42018

24 Layers

Transformer Encoder

Transformer Encoder

Transformer Encoder



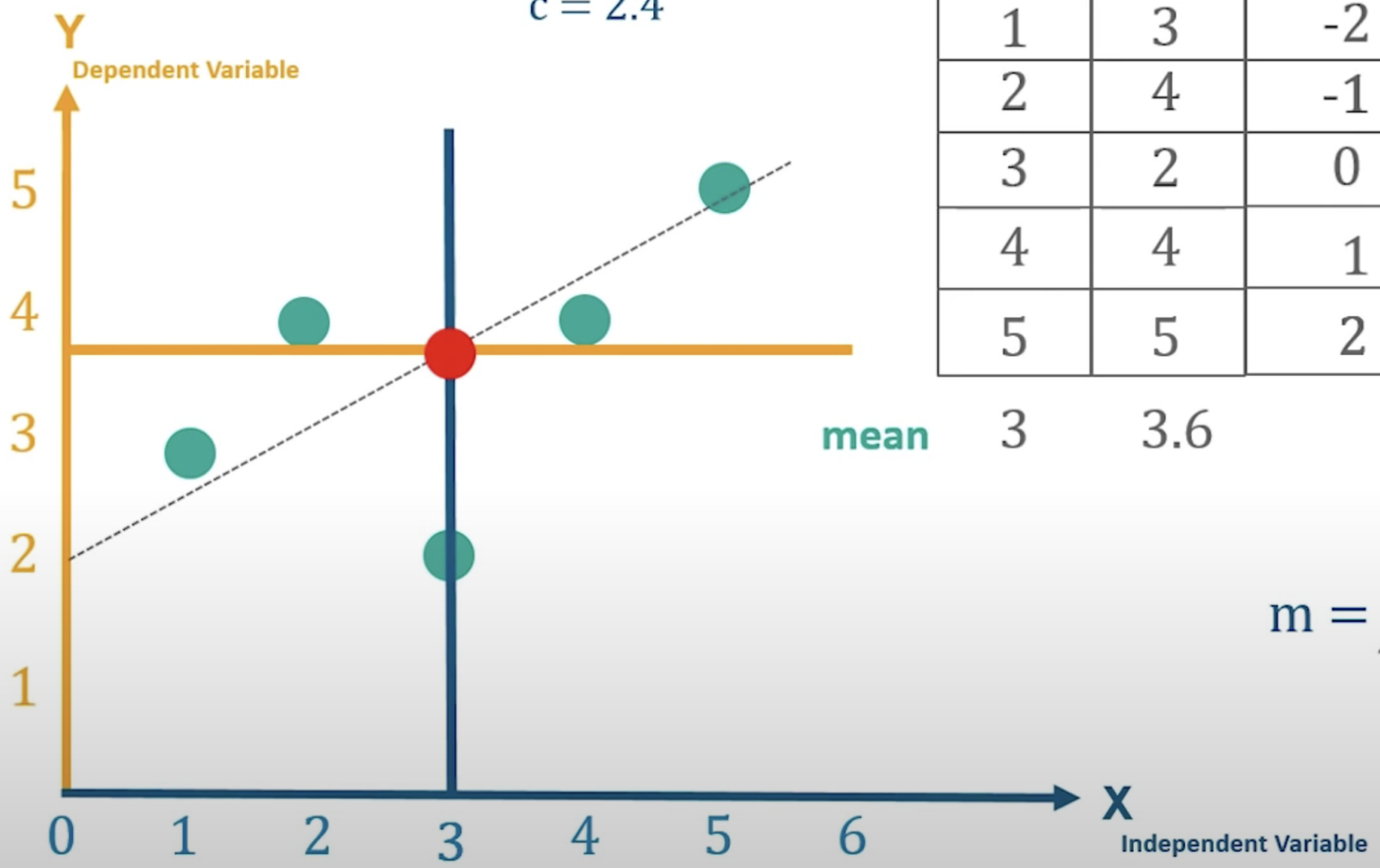
All ~~you need~~ is ~~Love~~
access to GPT-4 API

ML Recap

Understanding Linear Regression Algorithm

$$y = mx + c$$

$$c = 2.4$$



x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(x - \bar{x})(y - \bar{y})$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8
mean	3	3.6		$\Sigma = 10$	$\Sigma = 4$

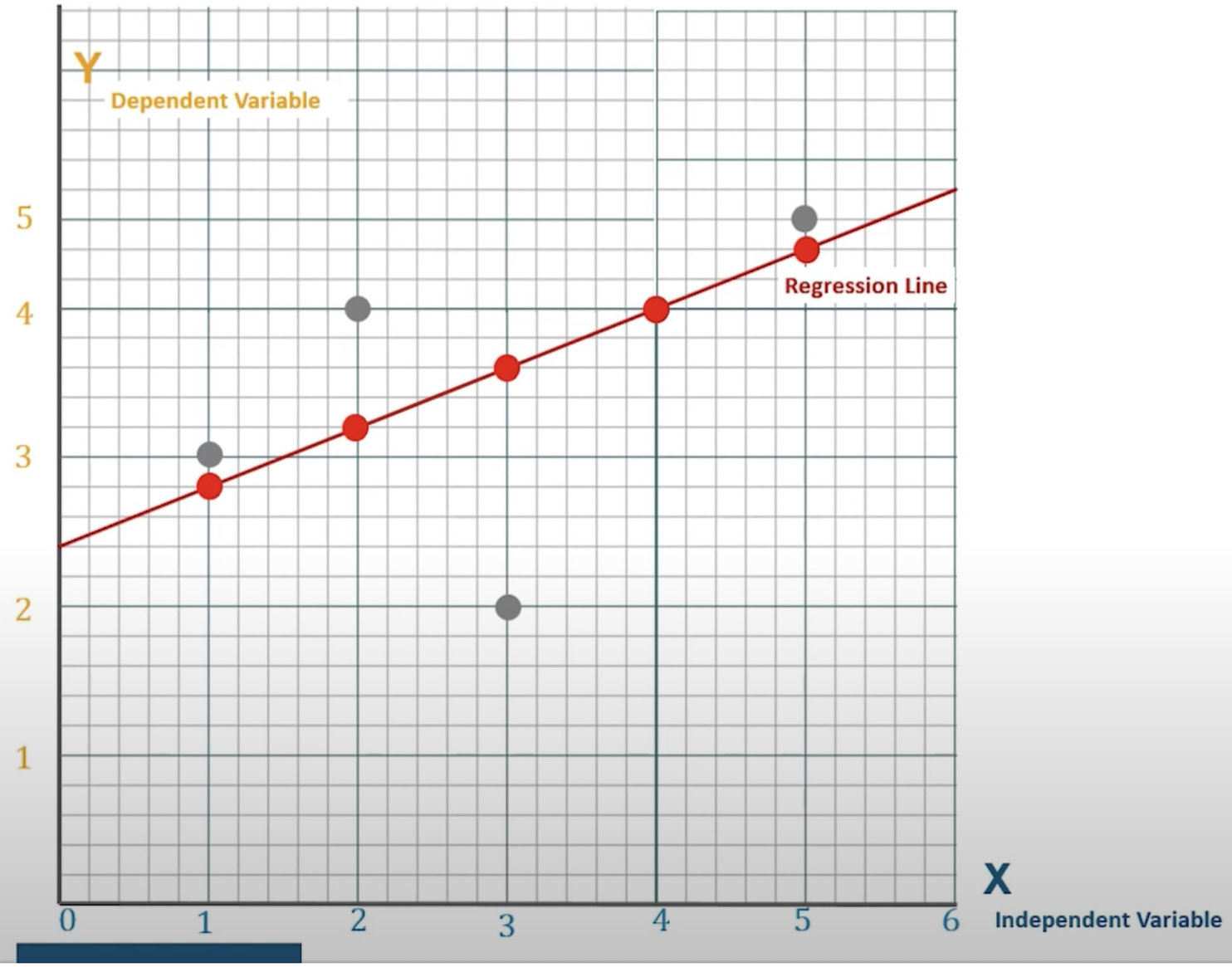
$$m = \sum \frac{(x - \bar{x})(y - \bar{y})}{(x - \bar{x})^2} = \frac{4}{10}$$

$$m = 0.4$$

$$c = 2.4$$

$$y = 0.4x + 2.4$$

Mean Square Error

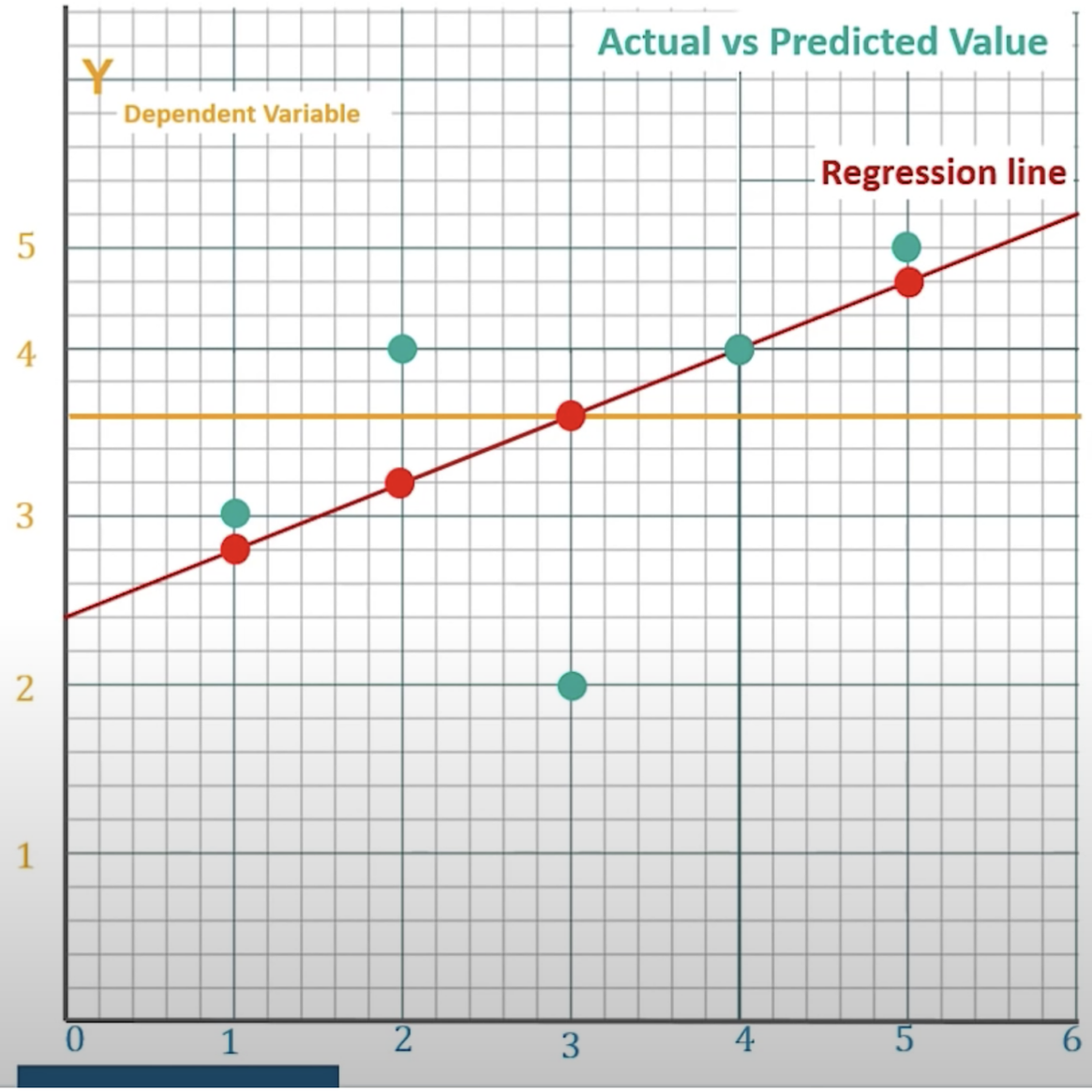


$$m = 0.4$$
$$c = 2.4$$
$$y = 0.4x + 2.4$$

For given $m = 0.4$ & $c = 2.4$, lets predict values for y for $x = \{1,2,3,4,5\}$

$$y = 0.4 \times 1 + 2.4 = 2.8$$
$$y = 0.4 \times 2 + 2.4 = 3.2$$
$$y = 0.4 \times 3 + 2.4 = 3.6$$
$$y = 0.4 \times 4 + 2.4 = 4.0$$
$$y = 0.4 \times 5 + 2.4 = 4.4$$

Calculation of R^2



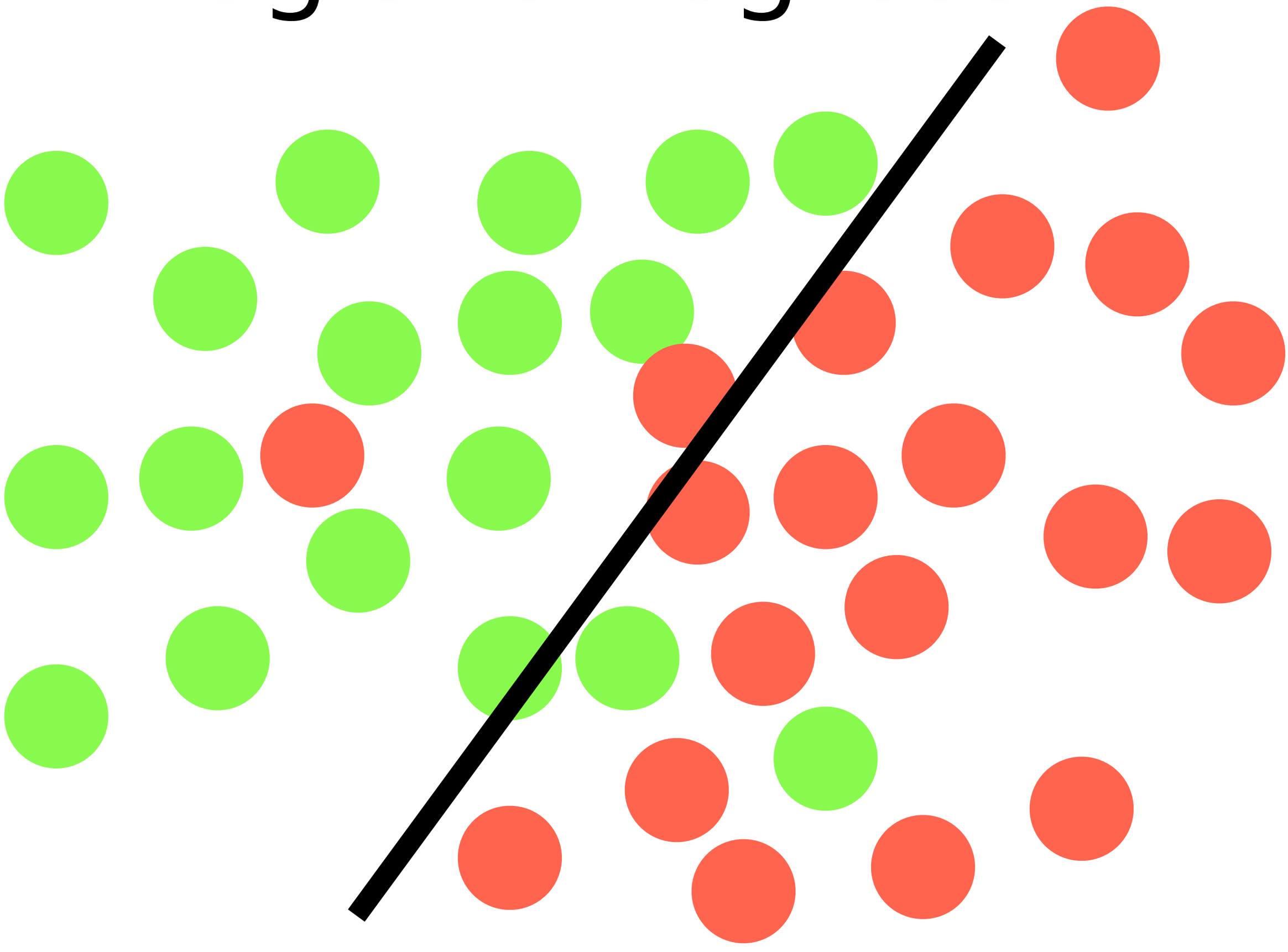
Distance actual - mean

vs

Distance predicted - mean

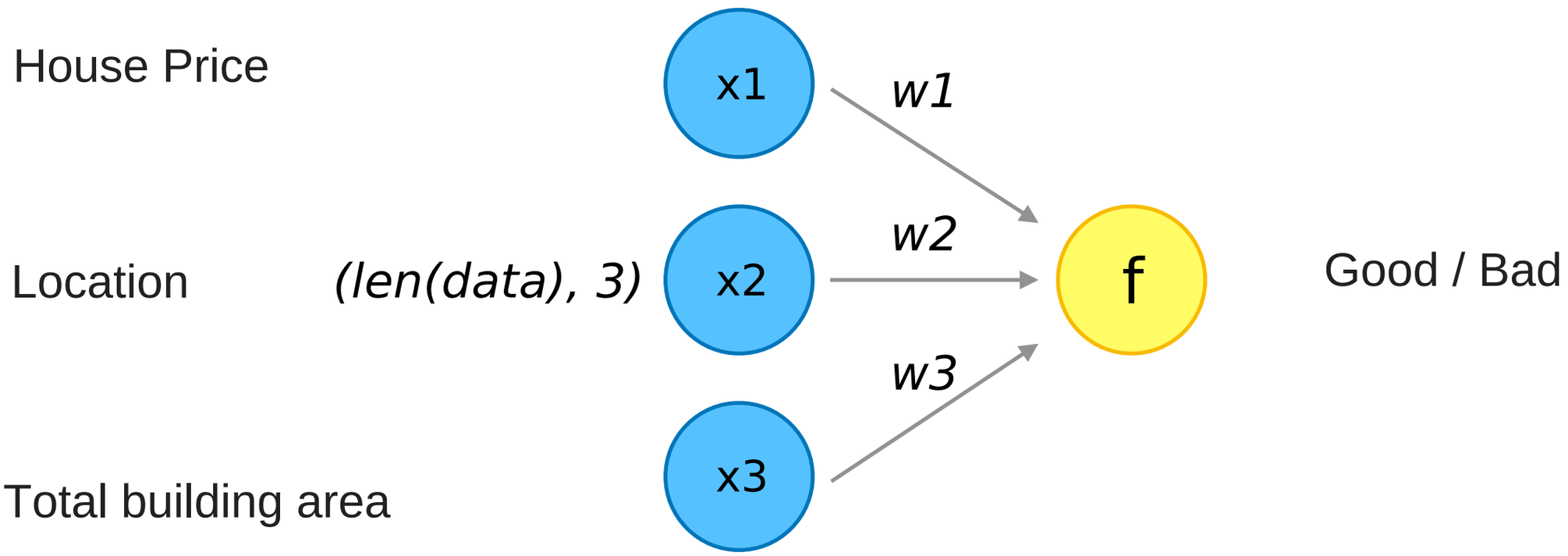
This is nothing but $R^2 = \frac{\sum (y_p - \bar{y})^2}{\sum (y - \bar{y})^2}$

Logistic Regression



Logistic Regression

Inference



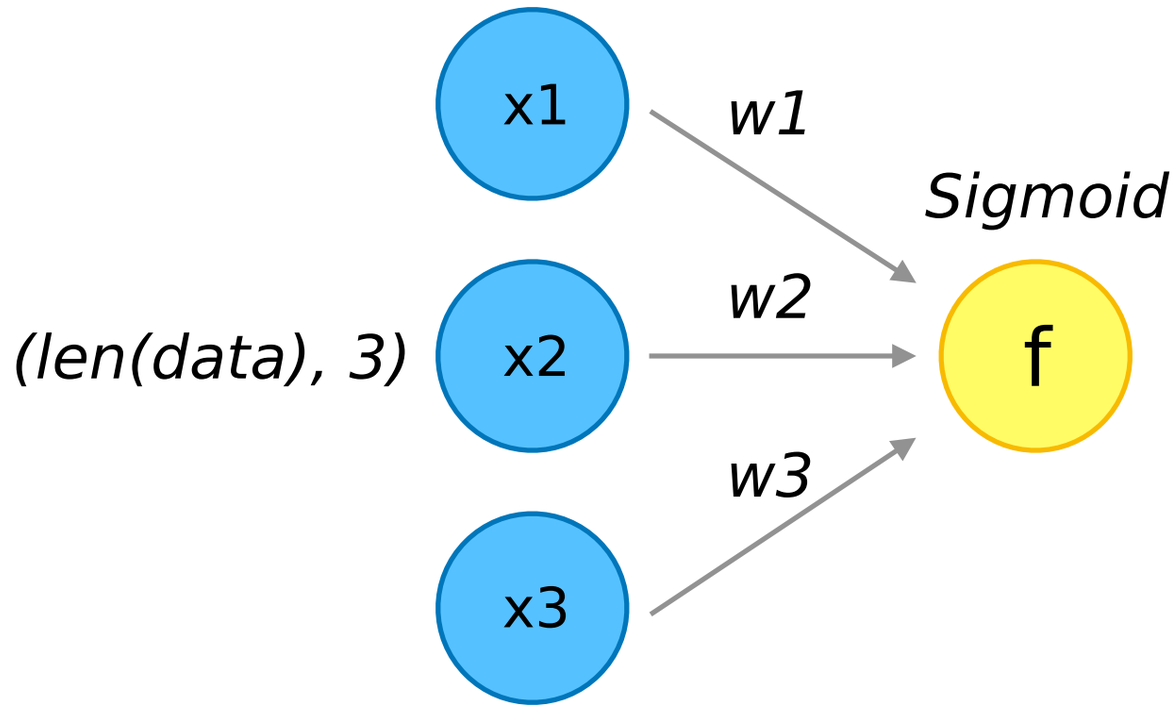
Logistic Regression

Inference

f = x * w + b

Logistic Regression

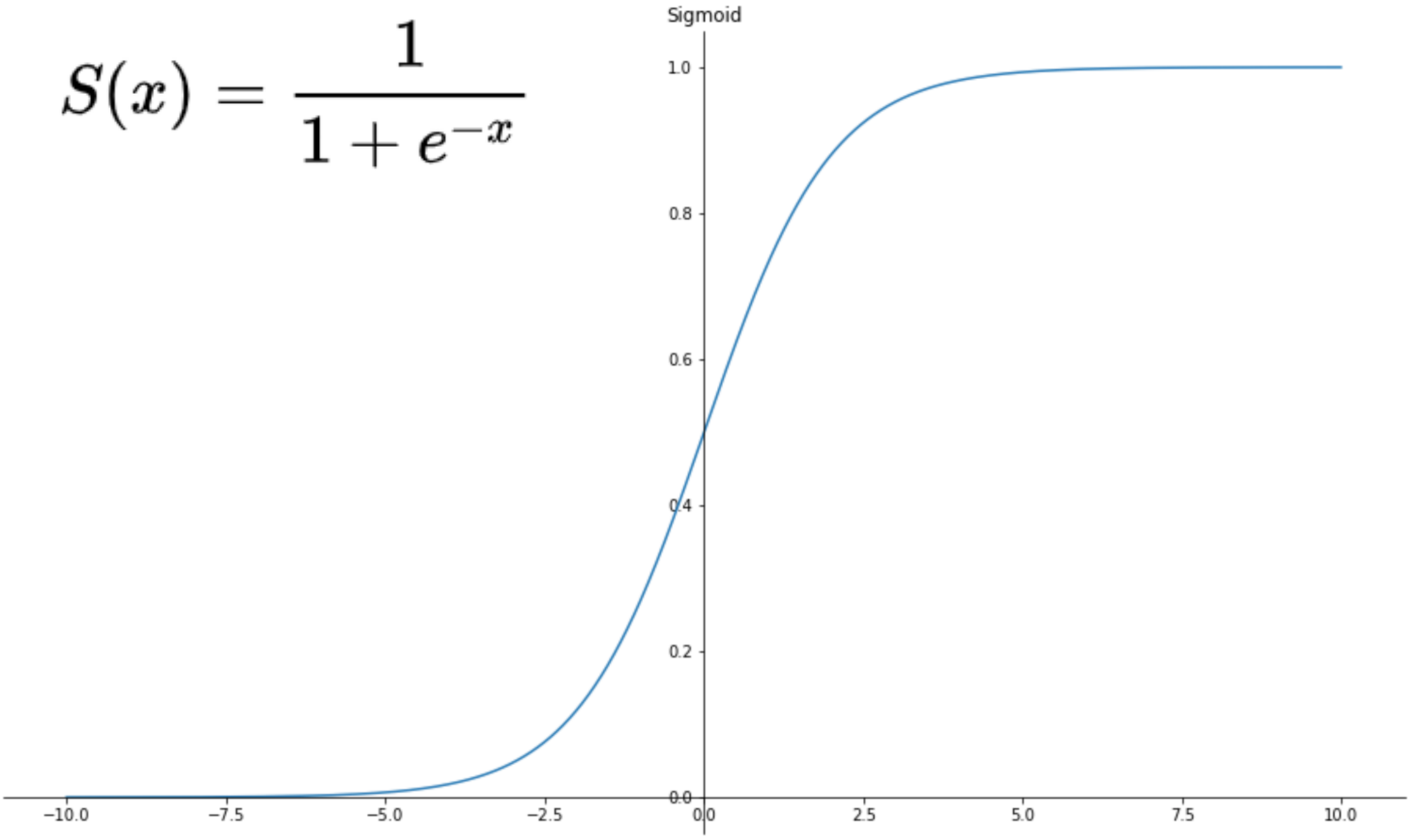
Inference



Logistic Regression

Sigmoid

$$S(x) = \frac{1}{1 + e^{-x}}$$



Logistic Regression

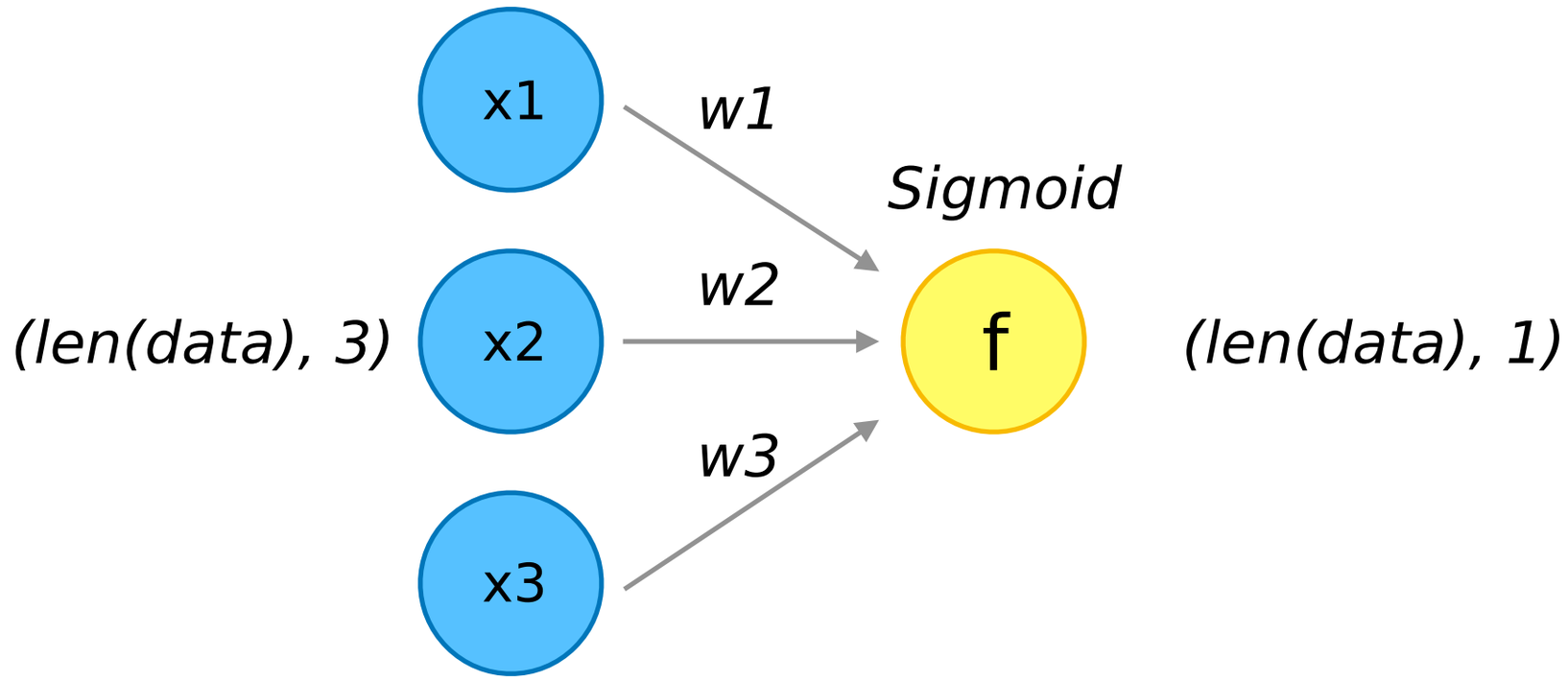
Inference

$$S(x) = \frac{1}{1 + e^{-x}}$$

f **= 1 / (1 + exp (- (x * w + b))**

Logistic Regression

Inference



w.shape = ?

Logistic Regression

Dot Product

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{s=1}^{n} a_{sn}b_{sj}$$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

f **= np.dot(x, w) + b**

$$AB = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \end{pmatrix}$$

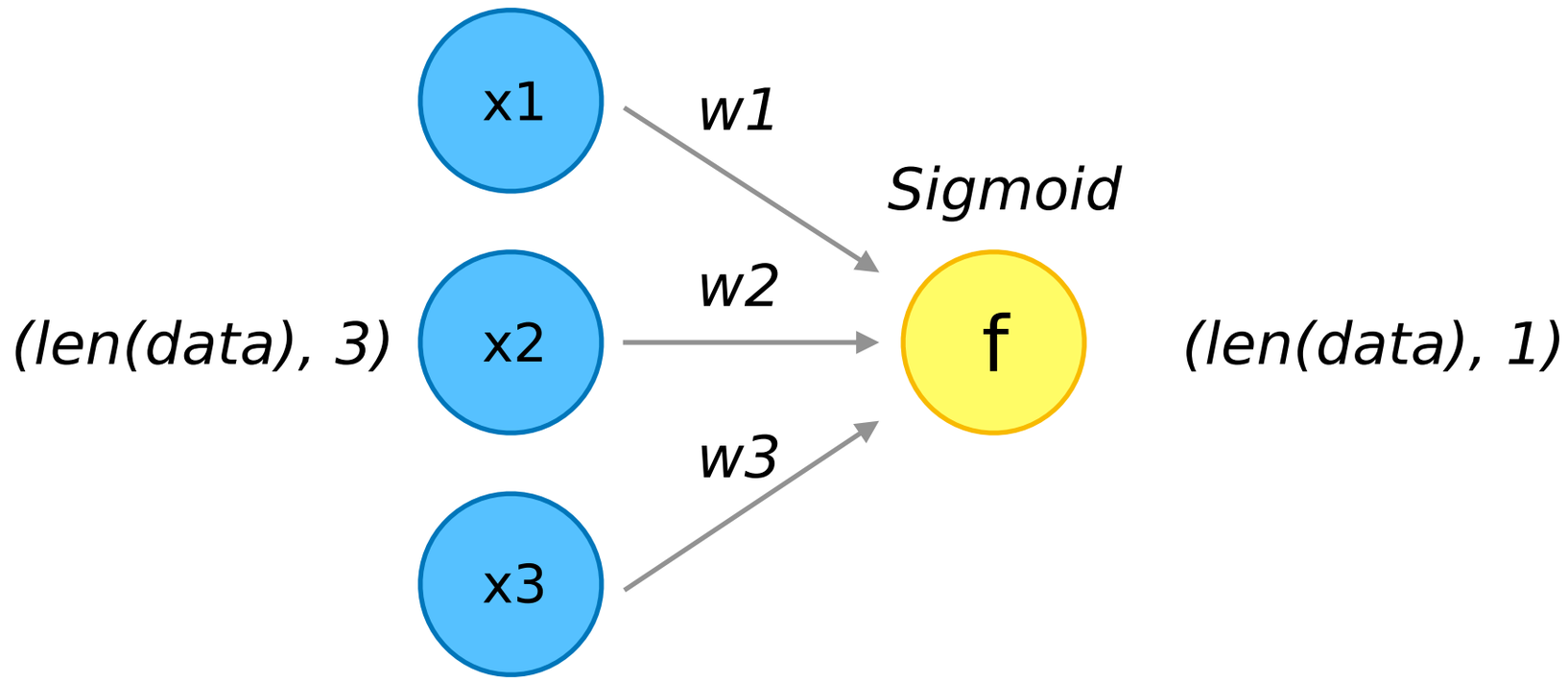
A.shape = (p, m)

B.shape = (n, k)

np.dot(A, B).shape = (p, k) if m =

Logistic Regression

Inference



w.shape = (3, 1)

Difference between classification and regression task

Classification predictive modeling problems are different from regression predictive modeling problems.

-- Classification is the task of predicting a discrete class label.

-- Regression is the task of predicting a continuous quantity.

There is some overlap between the algorithms for classification and regression; for example:

-- A classification algorithm may predict a continuous value, but the continuous value is in the form of a probability for a class label.

-- A regression algorithm may predict a discrete value, but the discrete value in the form of an integer quantity.

More info