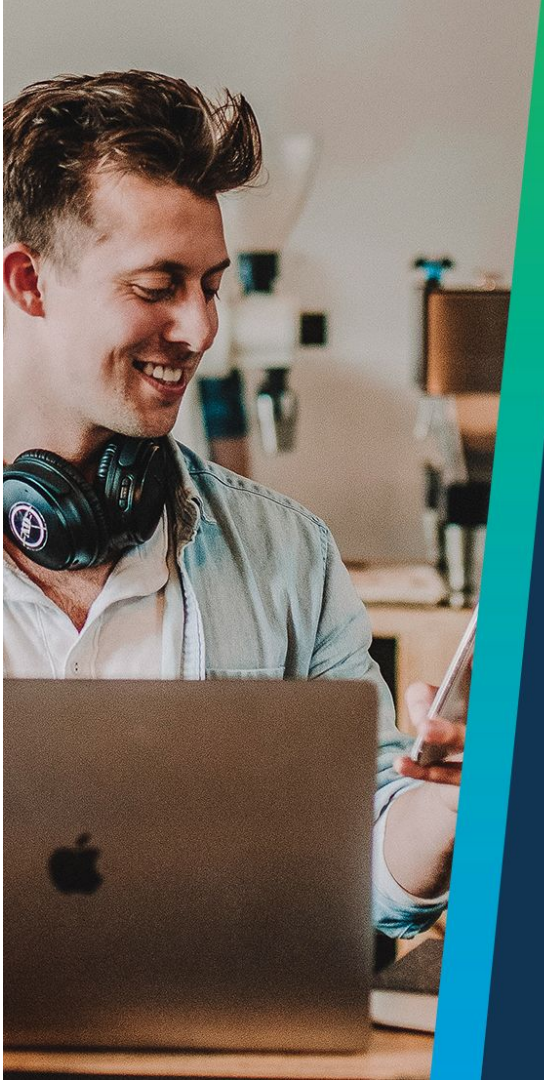


Data Science Bootcamp

Hyperiondev

Recap on Regression

Welcome

Your Lecturer for this session

Sanana Mwanawina

Lecture - Housekeeping

- The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly.
- No question is daft or silly - **ask them!**
- There are Q/A sessions midway and at the end of the session, should you wish to ask any follow-up questions.
- You can also submit questions here: hyperiondev.com/sbc4-ds-questions
- For all non-academic questions, please submit a query: hyperiondev.com/support
- Report a safeguarding incident: hyperiondev.com/safeguardreporting
- We would love your feedback on lectures: https://hyperionde.wufoo.com/forms/zsgv4m40ui4i0g/

Lecture - Code Repo

Go to: github.com/HyperionDevBootcamps

Then click on the "**C4_DS_lecture_examples**" repository, do view or download the code.

Objectives

- Solidify our understanding of regression by walking through a Logistic Regression problem

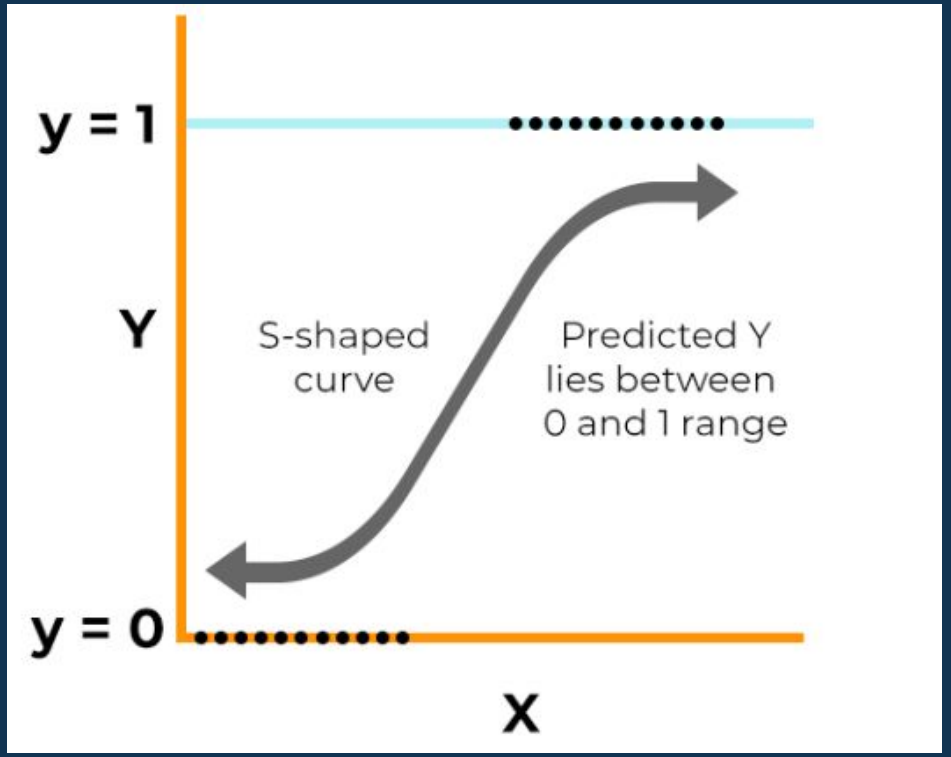
Our dataset

	1	2	3	4	5	6	7	8	9
1	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
2	6	148	72	35	0	33.6	0.627	50	1
3	1	85	66	29	0	26.6	0.351	31	0
4	8	183	64	0	0	23.3	0.672	32	1
5	1	89	66	23	94	28.1	0.167	21	0
6	0	137	40	35	168	43.1	2.288	33	1
7	5	116	74	0	0	25.6	0.201	30	0
8	3	78	50	32	88	31	0.248	26	1
9	10	115	0	0	0	35.3	0.134	29	0
10	2	197	70	45	543	30.5	0.158	53	1
11	8	125	96	0	0	0	0.232	54	1
12	4	110	92	0	0	37.6	0.191	30	0
13	10	168	74	0	0	38	0.537	34	1
14	10	139	80	0	0	27.1	1.441	57	0
15	1	189	60	23	846	30.1	0.398	59	1
16	5	166	72	19	175	25.8	0.587	51	1

Our problem

We have two patients. One with a glucose level of 120 and another with a glucose level of 220. We want to predict their respective probabilities of being diabetic and classify them accordingly.

Logistic Regression



$$P = \frac{e^{a+bX}}{1+e^{a+bX}}$$

$$P = \frac{1}{1+e^{-(a+bX)}}$$

Logistic Regression

$$P = \frac{\exp(-5.35+0.04x)}{1+\exp(-5.35+0.04x)}$$

For x = 120, $P = \frac{\exp(-5.35+0.04(120))}{1+\exp(-5.35+0.04(120))} = 0.3659$

For x = 220, $P = \frac{\exp(-5.35+0.04(220))}{1+\exp(-5.35+0.04(220))} = 0.969$

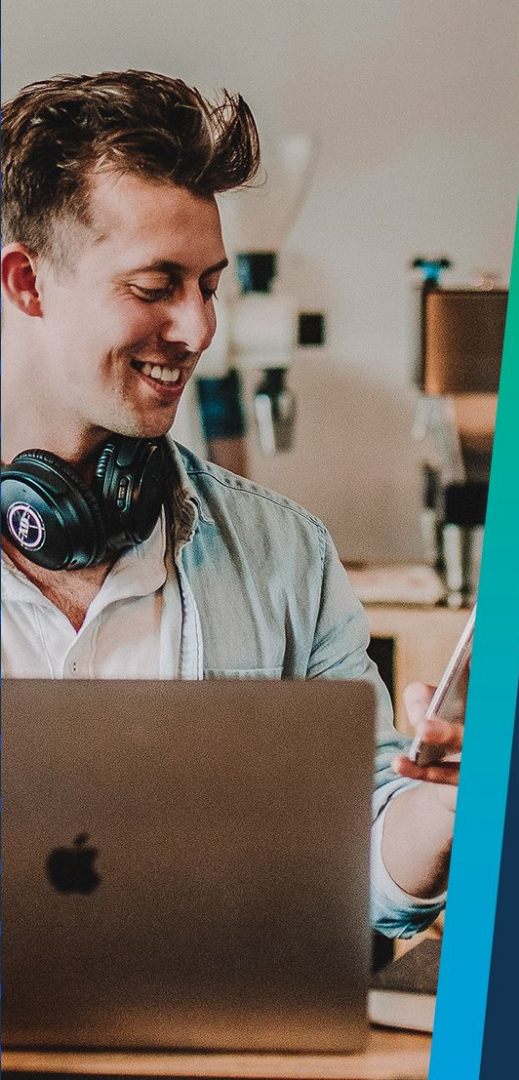
Multiple Logistic Regression

$$P = \frac{e^{(\beta + \alpha_1 X_1 + \ldots \alpha_n X_n)}}{1 + e^{(\beta + \alpha_1 X_1 + \ldots \alpha_n X_n)}}$$

Hyperiondev

Q & A Section

Please use this time to ask any questions relating to the topic explained, should you have any



Hyperiondev

Thank you for joining us