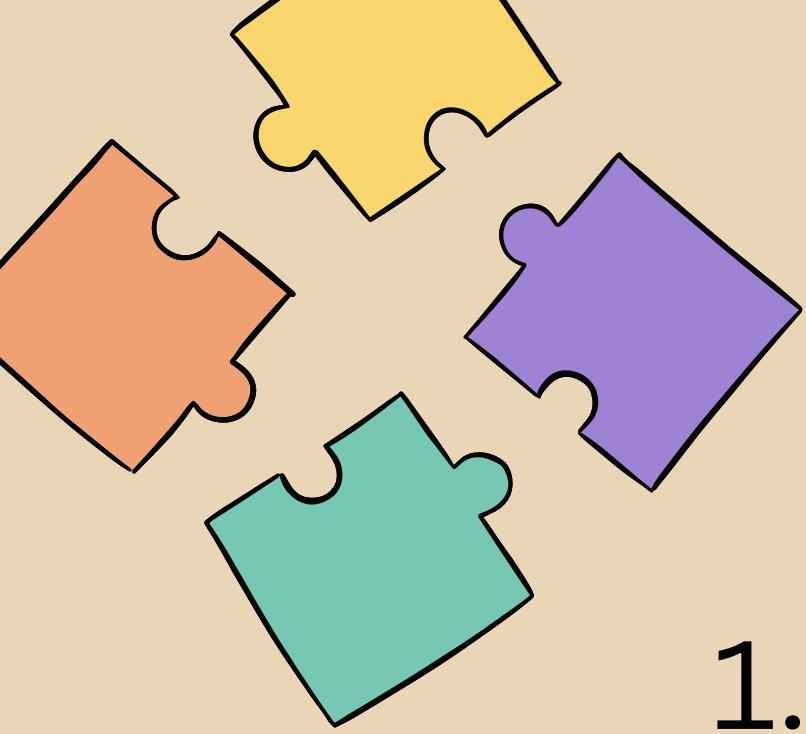


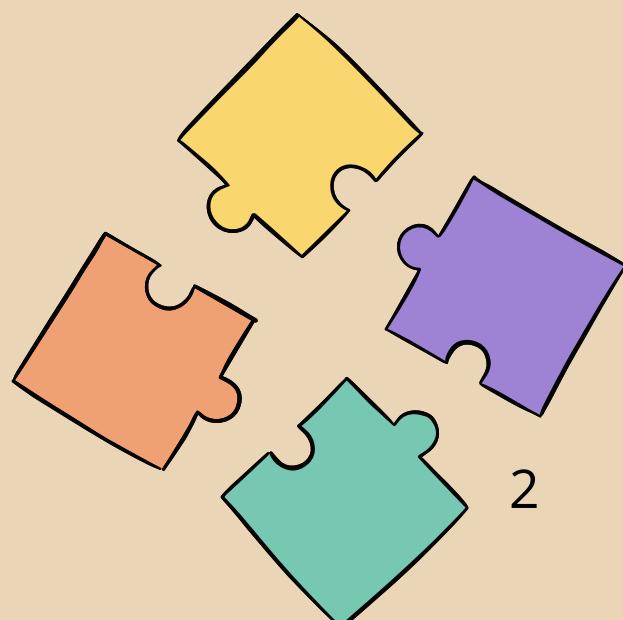
# LET'S BUILD OUR OWN MODEL



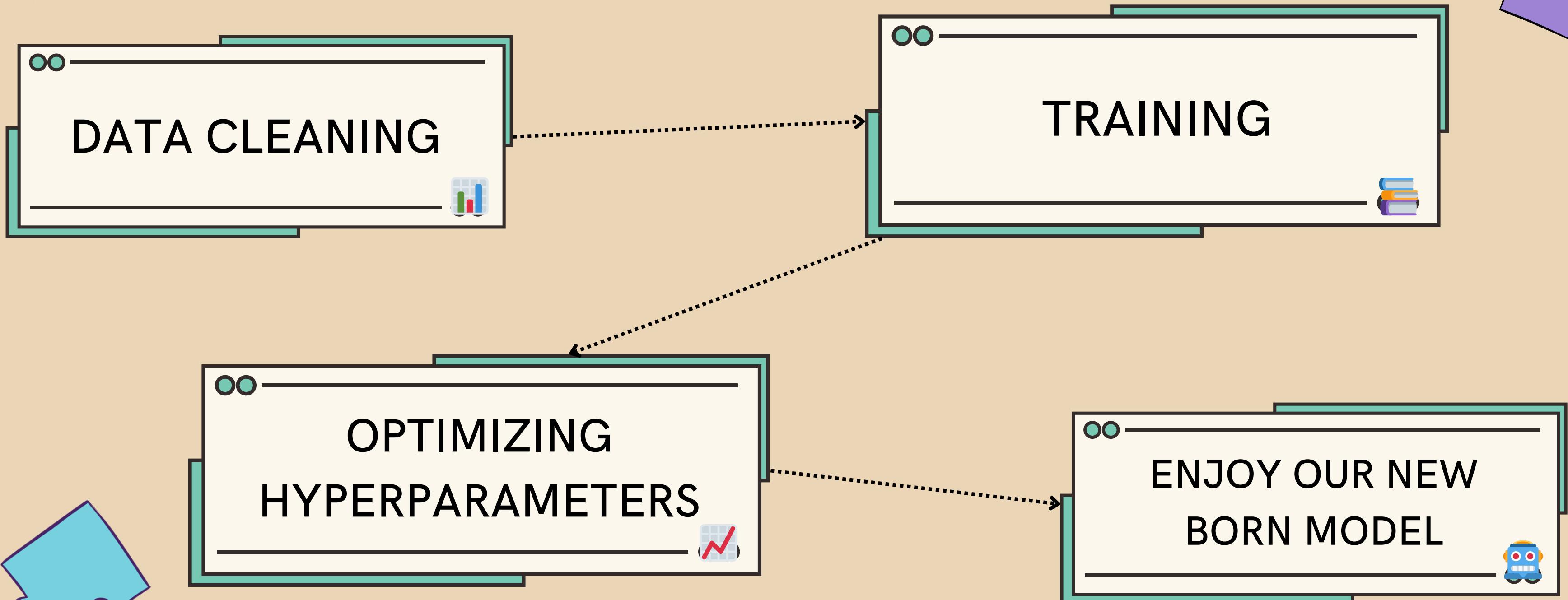


# Introduction

1. Workflow
2. Data
3. Model
4. Algorithm
5. Impact Study
6. Deliverables



# Our workflow



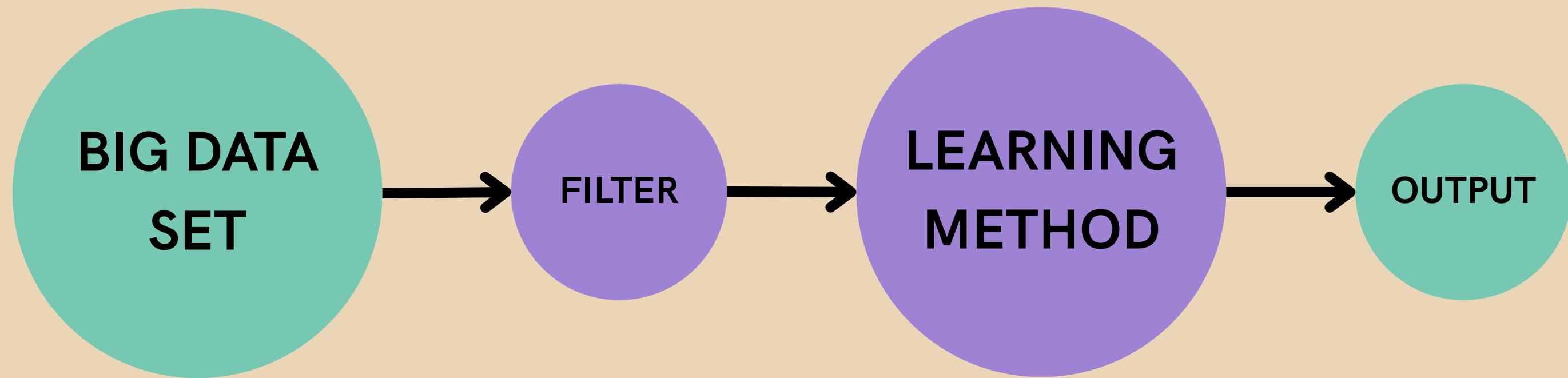
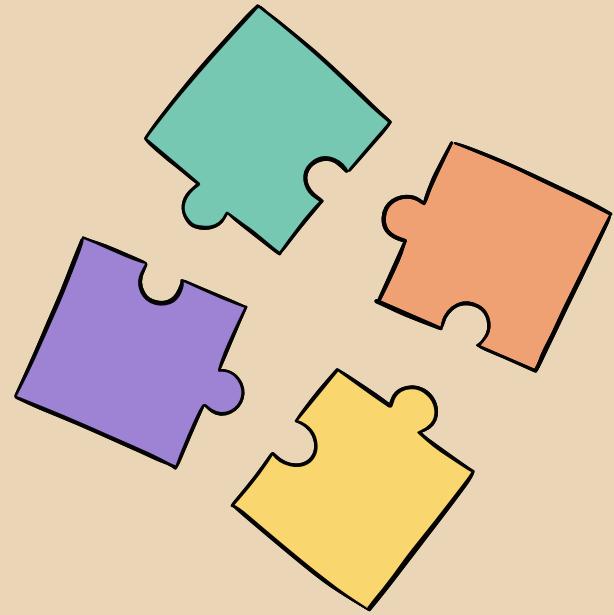
# DATA

- BIG Dataset —
- removing na values
- removing unique values (id, ...)
- removing duplicate columns
- removing date

ACHATS	AJUSTAGE	CN	ETUDES	METHODES	MOULE.CLEF.EN.M4_ST	Total_price	N1	N3	N10	N12	N18	Mail_Volume_Ad	Number_of_cavities	Nbre_mvt_mce_0	Nbre_mvt_mce_1	Nbre_mvt_mce_2	Nbre_mvt_mce_3	NB_Caviter	id	Mail_Kind	Injection_etape	A_Nbrr_total	unique_value_ID	V_D_Data	Capahegeli	Acier.moul.Regulation	Commande_Ejection	Type.Inject	Plastic_Part_APIA
521	0,10824557101663	0,2725095136211094	0,0772279701426192	0,2853615392524	0,24561362695491	0,16052893629732	0,167917205692109	0	0,25	0	0	0,04574910394261	0	0	0,1875	0	0	223	0	1	P1071221	06.10.2022	Mediatel1	Acier.moul.Chaud	Verin	Autre	BleuChaud.Thermaplastic.PPA.PPS.PEEK		
395	0,27072522563038	0,2688490336160612	0,21892343202852	0,0124556523211177	0,09933675007536	0,29798792755329	0,253186495065391	0,257438551969162	0	0	0	0,09643564885930	0,096774192548387	0	0	0	0	0,096774192548387	922	0	1	P18416401	06.11.2022	Mediatel1	Acier.moul.Chaud	Pezzo	Cylindrique	BleuChaud.Thermaplastic.PA.with.fiber	
690	0,220052425245209	0,34151842334548	0,25823756439488	0,50526949457404	0,32723767653937	0,222424274745	0,47484909457404	0,222424274745	0,25	0	0	0,12238690064647	0,032258064516129	0	0	0	0,16666666666667	0,032258064516129	440	0	3	P1135622	22.11.2022	Mediatel1	Acier.moul.Harmole	Pezzo	Cylindrique	BleuChaud.Thermaplastic.PA.with.fiber	
715	0,1763203220914426	0,1549016193527	0,04693008732995	0,090871264291974	0,21227364195107	0,104468789813621	0,222303836351876	0,25	0	0	0	0,05268463594470	0,096774192548387	0	0	0	0,096774192548387	412	0	1	P1132624	23.11.2022	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber		
555	0,0774264265195873	0,2593658757373	0,22471952349492	0,17923871264291974	0,26578006060590	0,175679172056921	0,00075	0	0	0	0	0,08523879501621	0	0	0	0,083333233323323	240	0	1	P1074822	18.11.2022	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PE.PP			
196	0,03232493473115916	0,24136243887616	0,169320362091928	0,2581624321303823	0,11589757954574	0,146183599870634	0,25	0	0	0	0,05075268817202	0,022580645161290	0	0	0	0,16666666666667	0,022580645161290	1017	0	1	P1843611	18.03.2020	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber		
248	0,0774264265195873	0,2593658757373	0,22471952349492	0,17923871264291974	0,26578006060590	0,175679172056921	0,00075	0	0	0	0	0,08523879501621	0	0	0	0,083333233323323	194	0	1	P1074715	18.03.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber			
141	0,2853924594591657	0,252705232352459	0,2413624321303823	0,17923871264291974	0,26578006060590	0,175679172056921	0,00025	0	0	0	0	0,0917530516972051	0	0	0	0,04166666666666	936	0	4	P1851711	19.07.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber			
670	0,02852050099924	0,091441447457426	0,17923871264291974	0,17923871264291974	0,0610640002094957	0,00025	0	0	0	0	0,0197530516972051	0	0	0,046774192548387	297	0	1	P1124521	24.10.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber					
230	0,10224794432327	0,25224045456000	0,050477226453493	0,1502564542080	0,1502564542080	0,050477226453493	0,25	0	0	0	0	0,02039423275587	0,032258064516129	0	0	0	0,096774192548387	1066	0	1	P1807113	13.12.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PE.PP		
180	0,190525141828	0,463061423898	0,2499819059439240	0,2499819059439240	0,2499819059439240	0,2499819059439240	0,25	0	0	0	0	0,207956428906451629	0,096774192548387	0	0	0	0,096774192548387	988	0	3	P1886211	15.07.2019	Mediatel1	Acier.moul.Harmole	Pezzo	Cylindrique	BleuChaud.Thermaplastic.PA.with.fiber		
400	0,2326191448354910	0,539848423192349	0,26228616052591	0,4423426296052591	0,4423426296052591	0,4423426296052591	0,25	0	0	0	0,208053133285202	0,032258064516129	0	0,125	0	0,04166666666666	31	0	1	P1016411	15.07.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.POM			
667	0,196303956715671	0,2593658757373	0,22471952349492	0,17923871264291974	0,26578006060590	0,175679172056921	0,00025	0	0	0	0,20267963816530	0,032258064516129	0,0425	0,10416666666667	451	0	1	P1141611	15.07.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber					
123	0,08308517179764	0,23290962911570	0,156507732158484	0,1451891527514781	0,1451891527514781	0,1451891527514781	0,25	0	0	0	0,02612558077779	0	0	0	0	0	0	0	1	P1804041	25.06.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PPA.PPS.PEEK			
75	0,047094002665190	0,302199166071739	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,25	0	0	0	0,042995273672033	0,096774192548387	0	0	0,0833333333333333	566	0	3	P1191713	11.06.2024	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber				
264	0,2451626245205910	0,302199166071739	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,25	0	0	0	0,042558070971905	0,032258064516129	0,125	0	0	0,0833333333333333	1158	0	3	P1930412	19.11.2020	Mediatel1	Acier.moul.Harmole	Verin	Autre	Carbone.LSR.avermaling			
248	0,230741728117777	0,23290962911570	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,25	0	0	0	0,04277114739477	0,0438770967741925	0	0	0,0833333333333333	107	0	1	P1812022	07.12.2022	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PPA.PPS.PEEK				
348	0,247074172811777	0,23290962911570	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,17923871264291974	0,25	0	0	0	0,04277114739477	0,0438770967741925	0	0	0,0833333333333333	1212	0	1	P1842452	07.12.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.PA.with.fiber				
269	0,145239434245207	0,410347912405287	0,2499819059439240	0,2499819059439240	0,2499819059439240	0,2499819059439240	0,25	0	0,12516425252763	0	0	0,09125	0,09125	0	0	0,0208333333333333	14	0	1	P1016411	25.05.2019	Mediatel1	Acier.moul.Harmole	Verin	Autre	BleuChaud.Thermaplastic.POM			
52	0,124349341412570	0,2374165592895	0,24164740405123	0,2374165592895	0,2374165592895	0,23741655																							

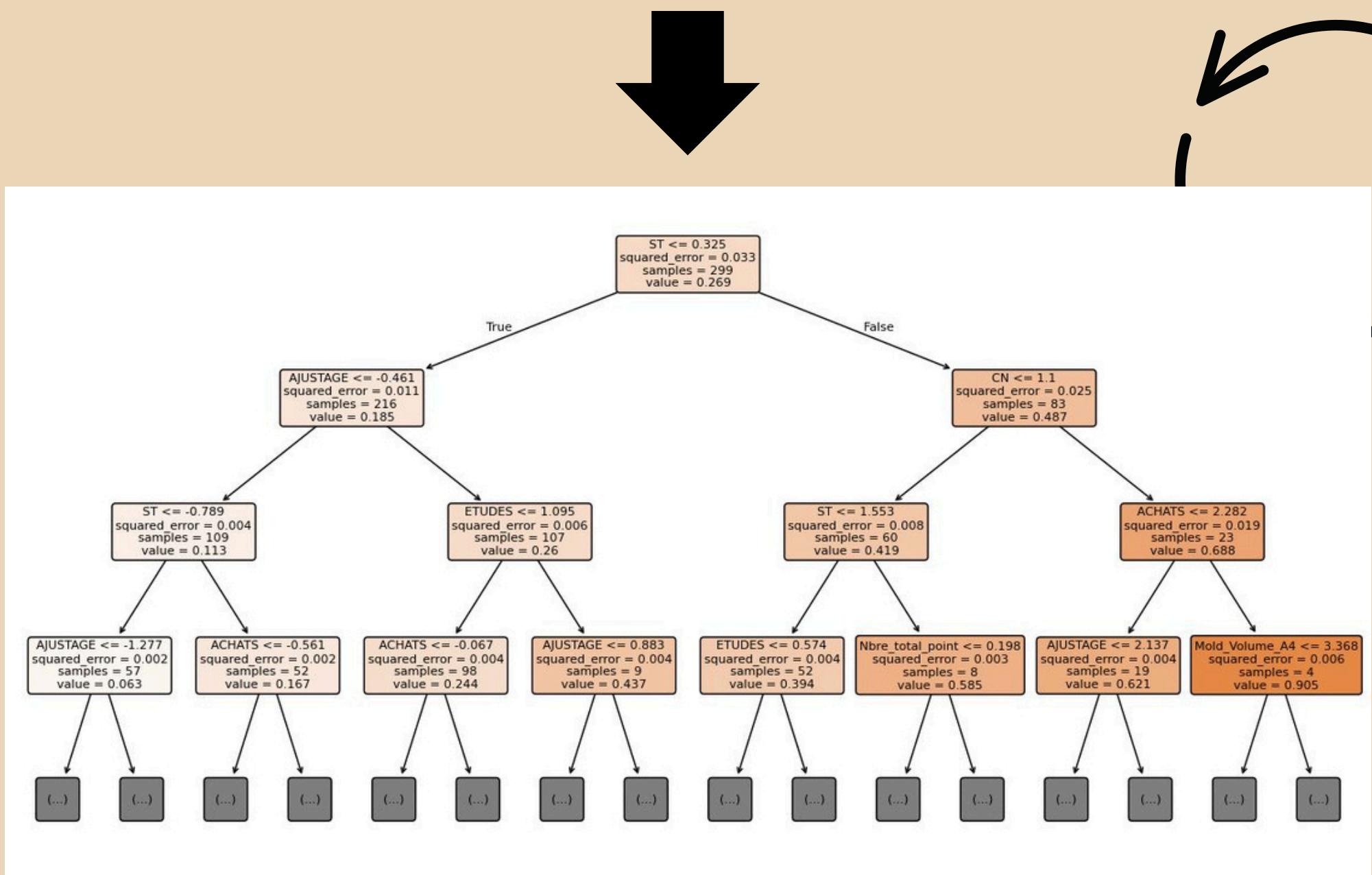


# MODEL



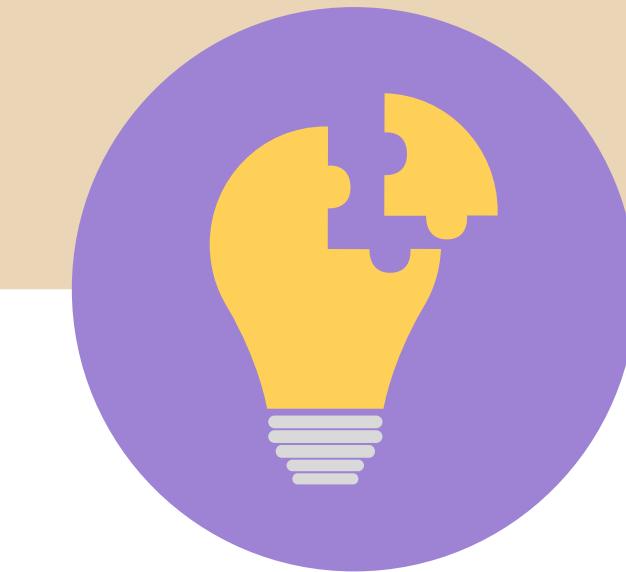
# ALGORITHM

## Random Forest vs. Gradient Boosting



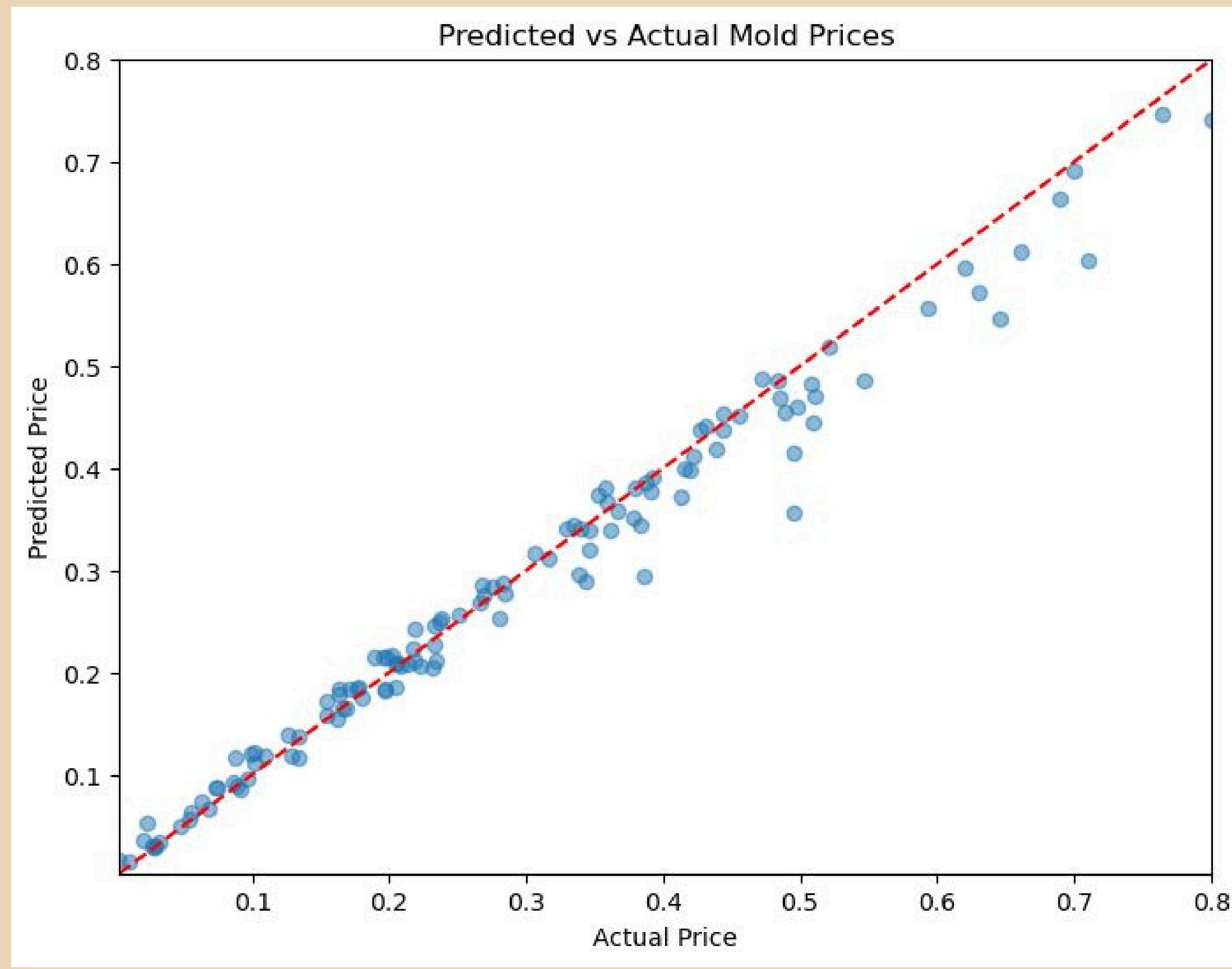
Hyper-parameter Tuning

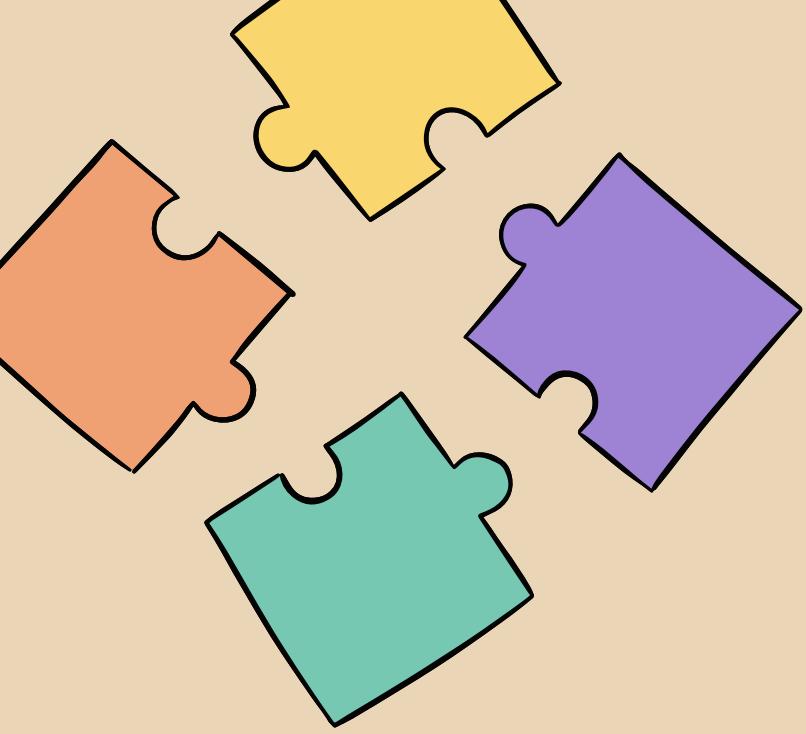
# And it works quite well



- ★ Mean Absolute Error (MAE): 0.02
- ★ Mean Absolute Percentage Error (MAPE): 0.10
- ★ Mean Squared Error (MSE): 0.00
- ★ Root Mean Squared Error (RMSE): 0.03
- ★ R<sup>2</sup> Score: 0.97

# And it works quite well



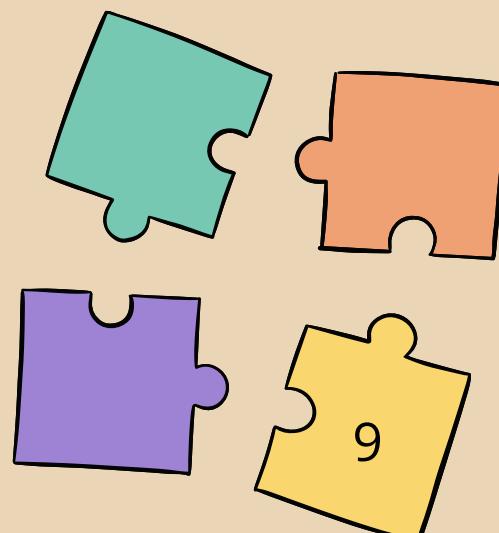


# **But it has an impact...**



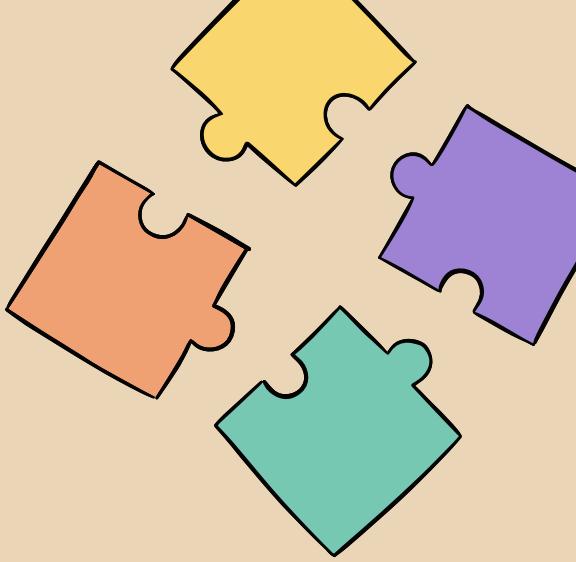
## **on the environment**

- 45 Seconds of Training on an Apple M1 Chip:
- 0.25–0.31 Wh
- €0.00005–€0.00010
- 7–8 mg of CO<sub>2</sub> → 1/5 of a single human exhalation





# Potential cyber risks



What if an attacker corrupts the data we use to manipulate our model and skew our predictions prices? This could impact our competitiveness!



What if our data leaks? And isn't anonymised? This could give lots of informations to our competitors and also violate GDPR laws



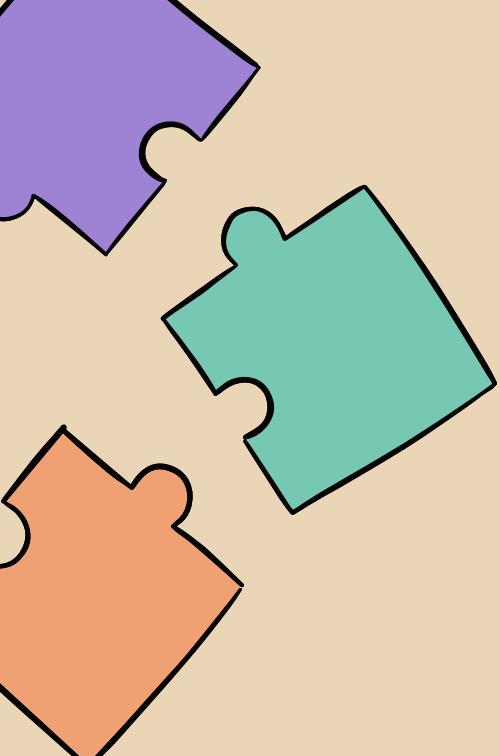
# **Change management**

## **Present:**

- Standardized template
- isolated functions (jupyter notebook)

## **Future:**

- Version Control System
- Testing & Validation Pipeline



# **And what about the maintenance?**

- Data Quality monitoring
- Model Performance tracking
- Infrastructure checks
- Retraining



**And Now...**

**It's Showtime**

# THANK YOU

