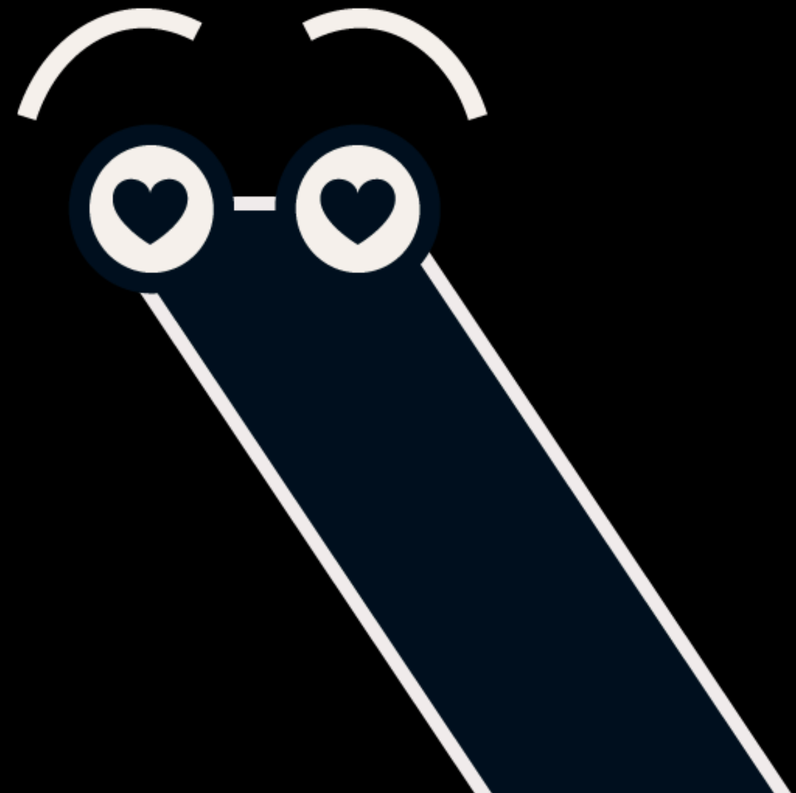
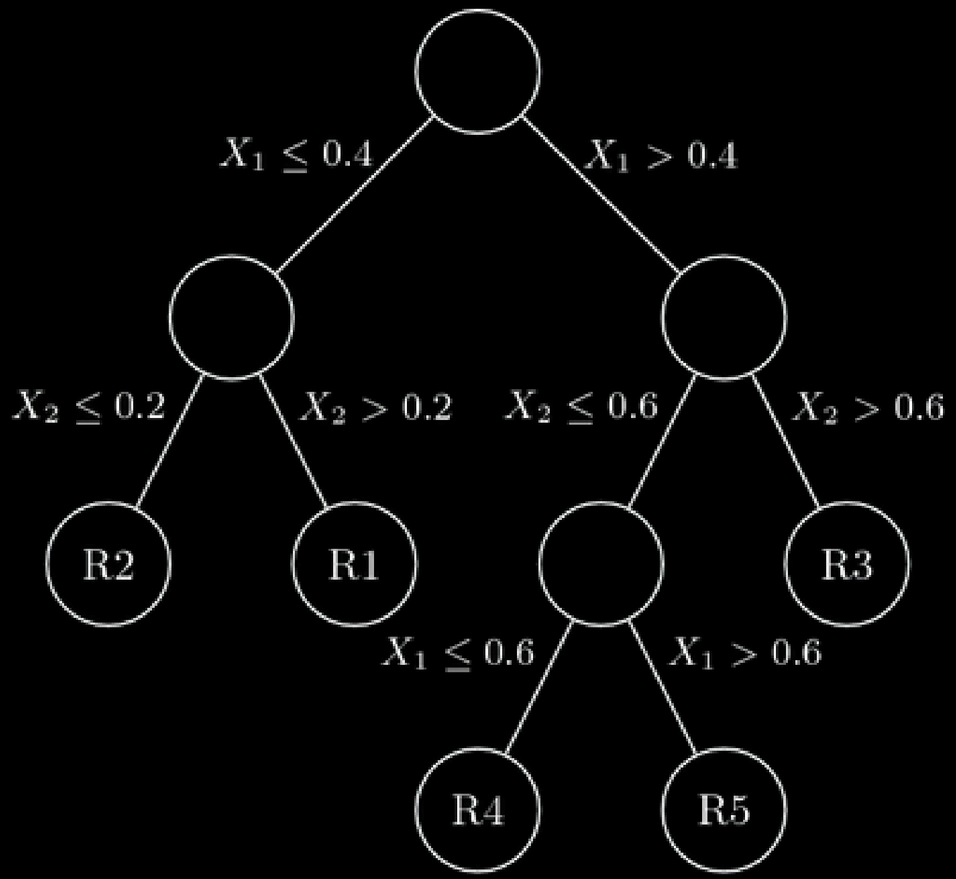
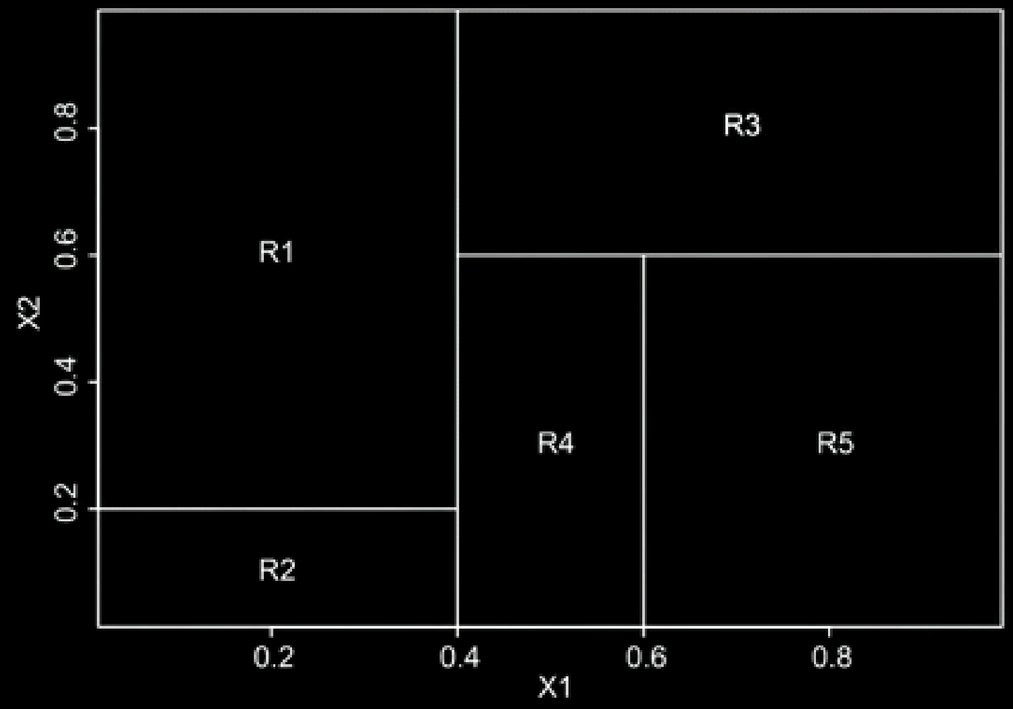
xtream

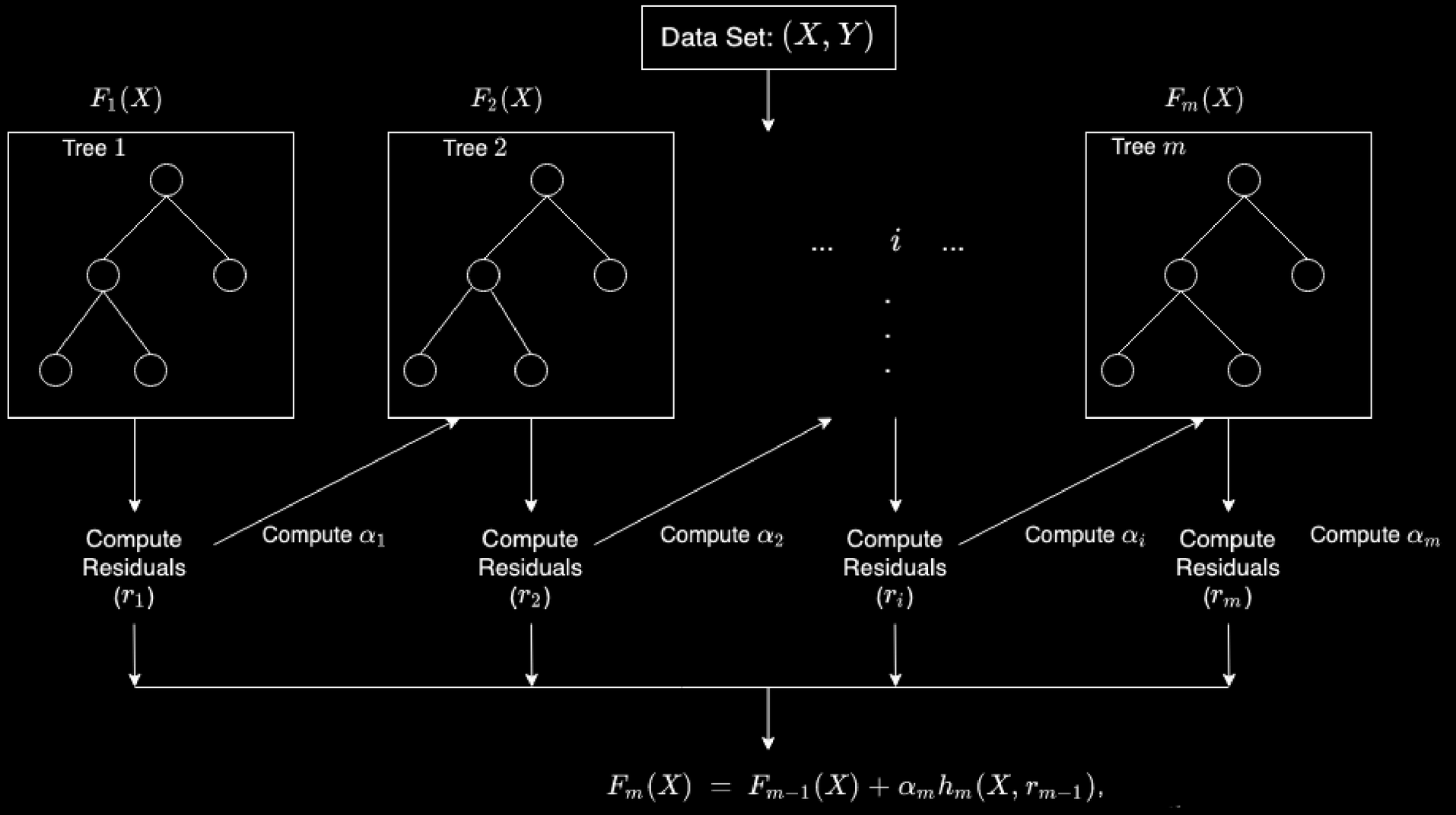
From SHAP to EBM

Emanuele Fabbiani




$X_1 \leq 0.4$
$X_1 > 0.4$
$X_2 \leq 0.2$
$X_2 > 0.2$
$X_2 \leq 0.6$
$X_2 > 0.6$
R2
R1
$X_1 \leq 0.6$
$X_1 > 0.6$
R3
R4
R5


R1
R3
R2
R4
R5
X2
0.2
0.4
0.6
0.8
X1
0.2
0.4
0.6
0.8


Data Set: (X, Y)
$F_1(X)$
$F_2(X)$
$F_m(X)$
Tree 1
Tree 2
Tree m
... i ...
Compute Residuals (r_1)
Compute α_1
Compute Residuals (r_2)
Compute α_2
Compute Residuals (r_i)
Compute α_i
Compute Residuals (r_m)
Compute α_m
$F_m(X) = F_{m-1}(X) + \alpha_m h_m(X, r_{m-1}),$

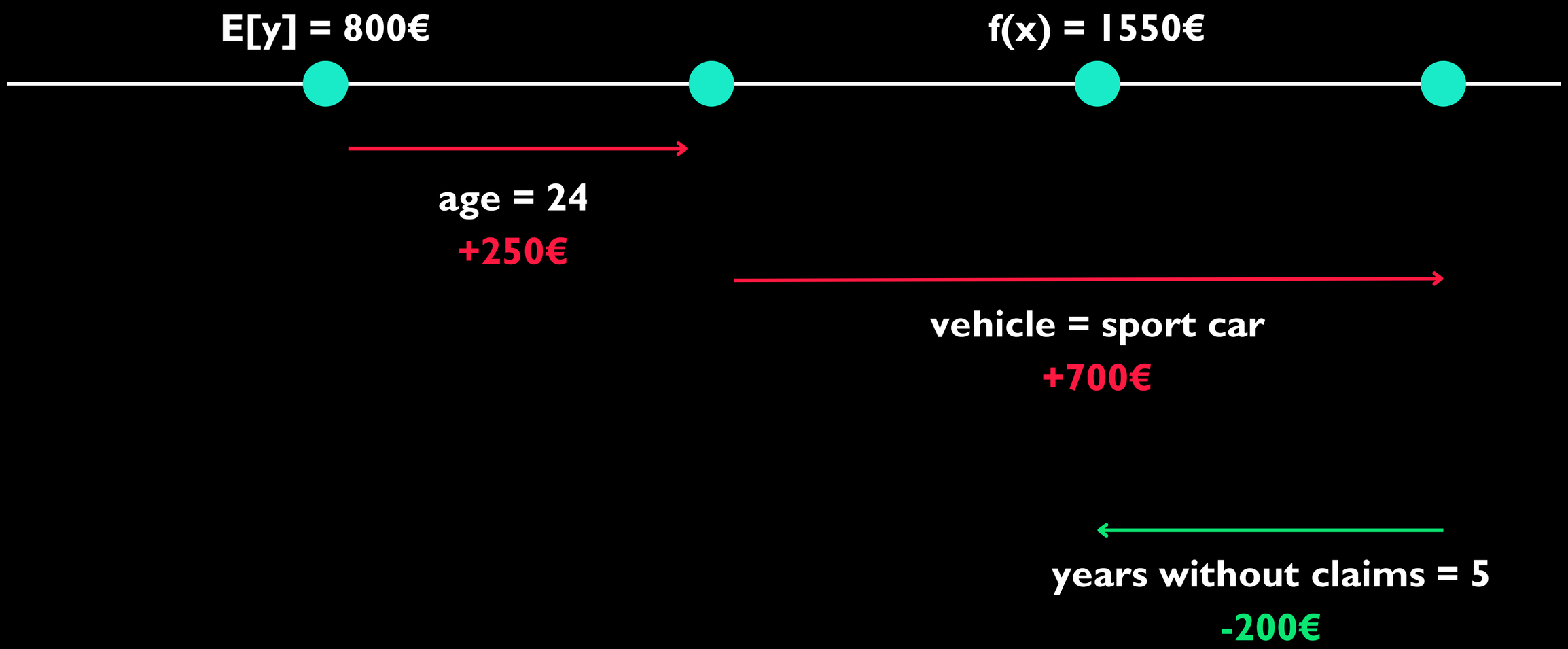
How can we measure the **impact** of each variable on the prediction for a **given sample**?

Local

Can be computed for each sample.

Additive

The effect of each feature sums up with the others.


E[y] = 800€
f(x) = 1550€
age = 24
+250€
vehicle = sport car
+700€
years without claims = 5
-200€

SHAP

Shapley Values

How do we fairly distribute money to players who win a game together, based on their individual **contributions**?

Efficiency

All the money must be distributed.

Consistency

If a player contributes more than another, they must get more money.

Theorem

There exists one and only one solution: split the money based on the **average** contribution of each player, over **all possible games** with every subset of players (in any order).

All possible combinations?

How to train the same model **without** some features?

All possible combinations?

Sample!

How to train the same model without some features?

Sample!

All possible combinations?

Sample!

(Efficient algorithms exist for tree and deep learning models.)

How to train the same model without some features?

Sample!

(From some dataset, with all kind of issues.)

SHAP

A collection of smart algorithms to approximate Shapley values.

That is, the only fair way of computing feature contribution is black-box models.

EBM

Iteration	Feature 1		Feature 2		Feature 3	
1	T(x1)	Residuals →	T(x2)	Residuals →	T(x3)	→
2	→ T(x1)	Residuals →	T(x2)	Residuals →	T(x3)	
...						

https://arxiv.org/pdf/1603.02754

https://xgboost.readthedocs.io/en/stable/tutorials/model.html

https://christophm.github.io/interpretable-ml-book/

https://arxiv.org/pdf/1705.07874

https://scikit-learn.org/1.5/modules/tree.html#decision-trees

https://youtu.be/-taOhqkiuIo?si=anl3HVeTi9N46Wbe

https://youtu.be/0yXtdkIL3Xk?si=5AKkl0i7NlU3qGmj

https://github.com/donlelef/shap-and-ebm-explain-your-gradient-boosting

The End

Ego Slide

Emanuele Fabbiani

AI Engineer @ xtream

Professor in AI at Catholic University of Milan

PhD in Applied AI

Speaker at AMLD Lausanne, ODSC London, PyData Berlin, PyCon Florence, PyData Milan, PyData Paris

Lecturer at UniPV, PoliMI, UniCatt, HSLU, Politechnika Wroclawska.