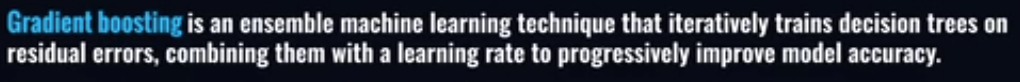
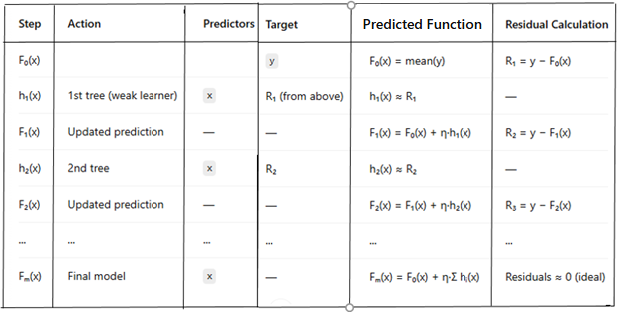
****

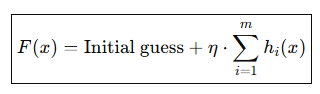
"**Gradient Boosting** = using gradient (errors) to boost your weak learners (individual decision trees)."

It’s like playing darts 🎯 — you keep throwing darts (trees), getting closer to the bullseye with each try!

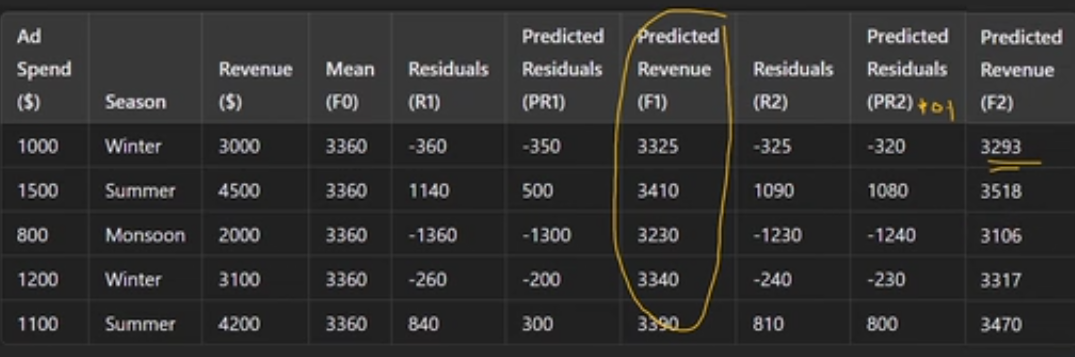


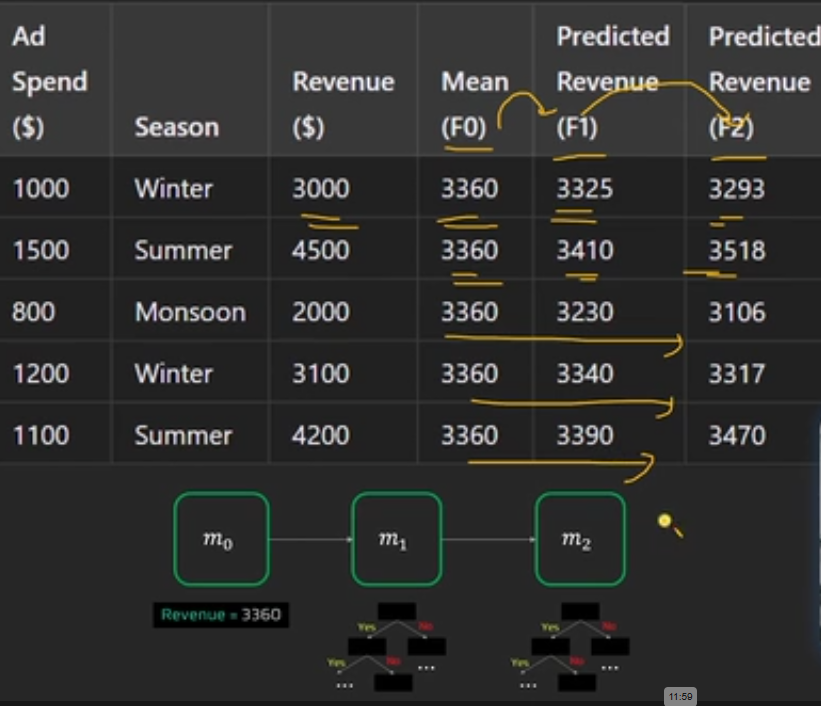
* h₁(x), h₂(x), ..., hₘ(x) are **hypothesis functions**, typically individual **small decision trees**, trained on the residuals from the previous predictions.
* η is the **learning rate**. (generally 0.1)
* The process continues for m steps or until the model converges (residuals become very small).
* The final model is **not one tree**, but a **sum of trees**, each fixing previous errors.
* Here F0​(x), F1​(x), …, Fₘ(x) are the **accumulated prediction** functions after adding corrections(boosting).

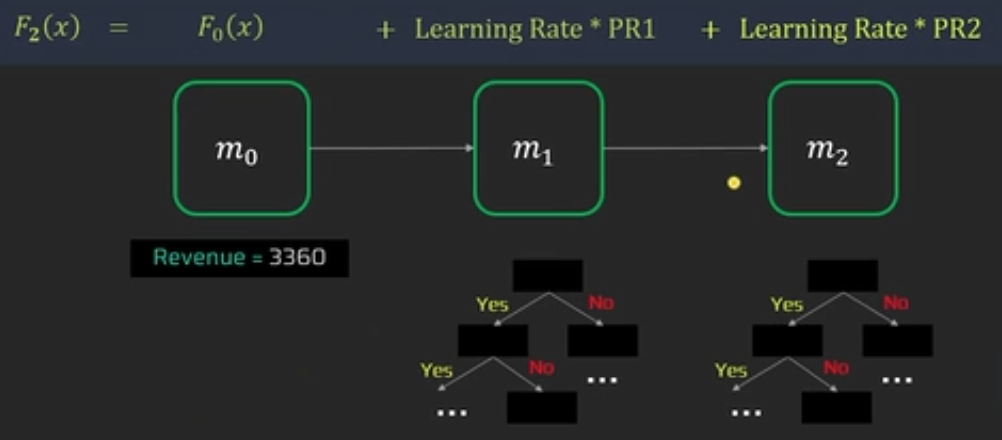
F = Initial guess + all tree corrections



**Example**

****

****

****