PPTs outline

1. Introduction to uplift model
2. Traditional models vs uplift model
3. What is uplift models different approach
4. An example of email marketing - using both model1 and model2 approach
5. Evaluation of both model - other than AUC method
6. Feature importance - Model explainability
7. Implemented case studies - An American bank summary

1. Introduction to Uplift Modeling and Its Objective

🔹 What Is Uplift Modeling?

Uplift modeling (also known as true lift modeling or incremental modeling) is a machine learning technique designed to estimate the **causal impact of an action** (like a marketing campaign) on individual customer behavior.

Unlike traditional response models that predict *who is likely to buy*, uplift modeling predicts:

> ***"Who is likely to buy because of the campaign?"***

🔹 **Why Is Uplift Important in Business?**

1. 📉 Avoid cannibalizing sales from customers who would've purchased anyway
2. 🛑 Prevent negative lift by avoiding customers annoyed by the campaign
3. 💸 Spend smarter: focus budget on high incremental impact
4. 🎯 Boost targeting precision: know not just *who will respond*, but *who needs a nudge*

2. Traditional models vs uplift model

Here's the complete content for slide 2: Traditional Models vs. Uplift Model, formatted clearly for a business audience:

🟦 2. Traditional Models vs. Uplift Model

🔹 **What Do Traditional Models Predict?**

Traditional response models (like logistic regression, decision trees, or random forests) predict:

“What is the probability this customer will buy?”

They do not distinguish between:

- **Customers who would buy regardless of the campaign**
- **Customers who buy only because of the campaign**
- **Customers who might even respond negatively to it**

These models are excellent at finding high-propensity buyers, but not incremental impact.

🔹 What Do Uplift Models Predict?

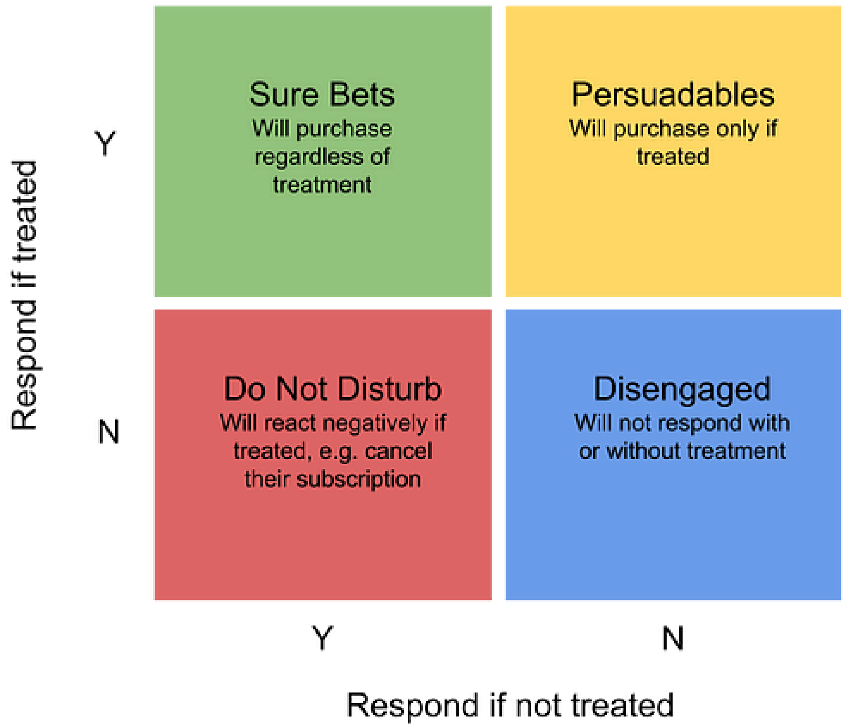
Uplift models ask a fundamentally different question:

“What is the difference in behavior *with* vs. *without* the treatment?”

They aim to measure the causal effect of the marketing action on each individual customer:

Uplift=P(Purchase | Treated)−P(Purchase | Not Treated)\text{Uplift} = P(\text{Purchase | Treated}) - P(\text{Purchase | Not Treated})

This gives you the true incremental gain from targeting a customer.



🔹 Key Differences: Table View

Aspect	Traditional Model	Uplift Model
Objective	Predict who will respond	Predict who will respond *because* of the campaign
Method	One model on historical responders	Compare treated vs. untreated populations
Bias Risk	May target customers who'd buy anyway	Focuses on persuadables
Controls	Not explicitly modeled	Requires control group
Targeting Outcome	High overall responders	High incremental responders (true lift)
Cost Efficiency	Lower	Higher ROI via better targeting

🔹 Example

Imagine these three customers:

Customer	Will Buy If Emailed	Will Buy If Not Emailed	Traditional Score	Uplift
A	✅	✅	High	0
B	✅	❌	High	1
C	❌	❌	Low	0

- Traditional model recommends both A and B
- Uplift model targets only B — the true persuadable

🧠 Final Takeaway

Traditional models optimize for response Uplift models optimize for influence — and that's what delivers ROI

3. Different Approaches in Uplift Modeling

🔷 a. One-Model Approach (S-Learner)

✅ Intuition

- Train one model using both treated and untreated customers.
- Add a special column: `treatment = 1` (if customer received campaign) or `0` (control).
- Also add interactions: multiply features by the treatment flag.
- At prediction time, simulate the customer with and without treatment to estimate uplift.

🧠 How It Works:

Let's say you use logistic regression:

$$P(\text{Purchase}) = \sigma\left(\beta_0 + \beta^\top X + \gamma \cdot T + \delta^\top (X \times T)\right)$$

- X: customer features (age, recency, spend history, etc.)
- T: treatment flag (1 = email sent)
- $\delta^\top (X \times T)$: how treatment effect varies across customer attributes

Then to get uplift:

$$\text{Uplift}(X) = P(\text{Purchase} \mid T = 1) - P(\text{Purchase} \mid T = 0)$$

💬 Pros:

- Simple implementation.
- Good when treatment effect is **smooth and consistent** across population.

⚠️ Cons:

- Can **underfit** if treatment and control behaviors are very different.
- Harder to interpret feature impact directly.

🔷 b. Two-Model Approach (T-Learner)

✅ Intuition

- Split your data into two groups:
  - Group 1: Treated customers
  - Group 2: Control customers
- Build **separate models** for each group, **both predicting probability of purchase**.
- For a new customer, get two predictions:
  - One assuming they **do get** the campaign.
  - One assuming they **don't**.
- The **difference** is your uplift estimate.

🧠 How It Works:

Let:

- $f_1(X)$: model trained on **treated** customers → predicts $P(\text{Purchase} \mid T = 1, X)$
- $f_0(X)$: model trained on **control** customers → predicts $P(\text{Purchase} \mid T = 0, X)$

Then:

$$\text{Uplift}(X) = f_1(X) - f_0(X)$$

💬 Pros:

- Flexible: each model can learn very different behavior.
- Can use **any ML algorithm** (XGBoost, GBM, etc.)

⚠️ Cons:

- May produce inconsistent probabilities (e.g., uplift > 1 or < 0).
- Requires good **data balance** between treated and control groups.

✅ Summary Table

Aspect	One-Model (S-Learner)	Two-Model (T-Learner)
# of Models	1	2
Treats T/C Differently	No (jointly modeled)	Yes (separate models)
Complexity	Lower	Higher
Flexibility	Moderate	High

Uplift Formula	f(X,T=1)−f(X,T=0)f(X, T=1) - f(X, T=0)	f1(X)−f0(X)f_1(X) - f_0(X)
When to Use	When treatment effect is consistent	When treated/control behave very differently

🟦 4. Example: Email Marketing Using Uplift Models >> Notebook is attached

🔹 a. Problem Statement

A retail company wants to increase sales by sending marketing emails to customers.
They run a campaign where:

- One group receives **Men's or Women's** promotional emails (treated group)
- Another group receives **no email** (control group)

Objective:

Use uplift modeling to identify customers who are likely to purchase **because of** the email (true lift), and avoid those who would purchase anyway or won't respond at all.

🔹 b. Introduction to Dataset

We use the **Hillstrom Email Campaign dataset**, a public dataset from a real marketing experiment.

Key contents:

- ~64,000 customers
- Treatment assignment: `segment` (Men's, Women's, or No Email)
- Behavior after campaign: visit and purchase
- Customer attributes: recency, history, channels, etc.

Feature	What It Means	Why It Matters

recency	“How many months has it been since this customer last bought anything?”	More recent buyers are generally more engaged—likely to respond again.
history	“How many total purchases did this customer make in the year before the campaign started?”	Heavy buyers (high history) often have different needs or loyalty than light buyers.
history_log	A “smoothed” version of history, taking the log of (history + 1) to tone down very large values.	Prevents extremely active customers from dominating the model by reducing skew.
history_log_squared	The square of `history_log`, capturing *non-linear* effects—e.g. very heavy buyers might behave differently from moderate ones.	Lets the model detect if there’s a special “super-buyer” segment whose reaction to emails is unique.
channel	“Which way does this customer usually buy: online (Web), by mail order, or both?”	Different channels imply different habits—an online shopper may respond better to an email than a mail-order

		shopper, for example.
segment	“Which version of the email did they get? Men’s email, Women’s email, or no email at all?”	Tells us who was in the treated group vs. the hold-out control group.
treated	A simplified flag: “Did this customer get *any* marketing email (1 = yes, 0 = no)?”	Transforms the three-way segment into a straightforward “treatment vs. control” indicator for uplift modeling.
mens	“Did this customer get the Men’s version of the email? (1 = yes, 0 = no)”	Allows analysis of whether the men’s email specifically drives lift, separate from the women’s email.
womens	“Did this customer get the Women’s version of the email? (1 = yes, 0 = no)”	Same as above, but for the women’s email.
newbie	“Has this customer *never* bought anything before?” (1 = yes, 0 = no)	First-time buyers can be more sensitive to marketing—may

		need different messaging.
visit_2wk	“Did the customer visit the website within two weeks after the email?” (1 = yes, 0 = no)	A quick sign of engagement—ofte n a precursor to spending.
sales_2wk	“How much money (in dollars) did this customer spend in the two weeks after the email?”	The *actual* business metric we care about—revenue generated by the campaign.
responded (alias: **conversio**)	“Did the customer either visit *or* spend money in those two weeks?” (1 = yes, 0 = no)	A simplified “yes/no” label combining visits and purchases, used as the target for both propensity and uplift models.

How These Features Address the Problem

1. recency & history tell us *who* is already active—traditional models would aim emails at your most active buyers, but uplift models want to know *who needs* the email to respond.

2. channel, mens, womens help us customize which message or channel is most persuasive.

3. newbie flags first-timers—sometimes a special welcome message may lift them more.

4. treated vs responded (or conversion) allows uplift modeling to compare what happened *with* vs *without* the email, isolating the true lift from each feature combination.

🎯 5. Evaluation of Uplift Models (Other Than AUC)

Uplift modeling has **special evaluation metrics** that go beyond AUC (which is suited for classification models like propensity). Here's how:

🅰️ a. Different Evaluation Methods for Uplift Models

Metric	Used In	Interpretation
Qini Score	Uplift models	Measures **incremental gain** over random targeting. Analogous to AUC.
Uplift@k%	Uplift models	Measures the difference in response rates between treated vs control in top k%.
AUUC (Area under Uplift Curve)	Uplift models	Similar to Qini, but less adjusted for randomness.
Decile-wise uplift	Uplift models	Difference in conversion between treated and control in each decile of predictions.
Persuadability lift	Advanced uplift	Focuses on identifying only **persuadable** customers.

📘 Why AUC Isn't Enough

- AUC works for **binary classification** — predicts **probability of response**.
- Uplift modeling is **causal** — it predicts the **difference in probability with vs without treatment**, i.e., *incremental impact*.

➡️ So we need metrics that measure **impact**, not just likelihood.

📊 b. Decile-wise Uplift (Decile Chart)

1. **Sort customers** by predicted uplift descending
2. **Split** into 10 equal-sized groups (deciles)
3. **For each decile:**
    - Calculate average response of **treated** group
    - Calculate average response of **control** group
    - Compute `Uplift = Treated_Response - Control_Response`

🔍 **Interpretation**:

- **Top decile** should show the **highest uplift** — those are your most persuadable customers.
- A decile chart helps visualize **where to target** for maximum ROI.

📐 What is Qini Score?

🧠 Definition:

Qini Score = **Area between uplift curve and random targeting line**

It's a generalization of AUC for uplift.

🎨 Analogy with AUC:

AUC (ROC)	Qini Curve
True Positive Rate vs False Positive Rate	Cumulative Response (Treated vs Control)
Area under ROC curve = performance	Area between Qini and random = uplift gain
Random baseline = diagonal	Random baseline = diagonal uplift

➡️ **Higher Qini score = better identification of impactable users**

🧮 Qini Curve Steps:

1. **Sort by predicted uplift**
2. **Plot cumulative response in treated group**
3. **Subtract cumulative response in control group**
4. **Plot difference as Qini curve**
5. **Qini score = area under this curve**

✅ Summary Table

Metric	Purpose	Uplift-Specific?	Analogy
Qini Score	**Overall model performance (incremental)**	✅ **Yes**	**AUC (for uplift)**
Uplift@K	**Precision targeting performance**	✅ **Yes**	**Precision@K**

Decile Uplift	Insight into segmentation performance	✅ Yes	Gains chart
AUUC	Area under uplift curve	✅ Yes	AUC variant
AUC	Traditional response model accuracy	❌ No	ROC Curve

6. Feature importance - Model explainability >>To be added in code notebook

7. 📊 Uplift Modeling at U.S. Bank case study: Key Takeaways

🧭 Challenge: Traditional Campaigns Lost Direction
U.S. Bank noticed that many customers were responding to offers even without receiving them — like watering a plant that's already thriving. Traditional models overestimated marketing impact.

🎯 Goal: Measure True Incremental Impact
They shifted focus from total responses to incremental responses — i.e., how much of the success was *caused* by the marketing, not just correlated with it. That's where uplift modeling came in.

🛠️ Solution: Adopted Portrait Uplift Modeling
After experimenting with internal models that didn't generalize well, U.S. Bank chose Portrait Uplift, which helped them robustly predict the difference in response between treated and control groups.

📈 Massive Business Gains
Like giving coupons only to those who need a nudge, uplift modeling helped:

- **🔼 Boost cross-sell revenue by over 300%**
- **💰 Add $1 million+ in incremental revenue from just 2 campaigns**
- **📬 Reduce mailing volume by up to 40%, saving costs**

🧪 Test and Learn Culture
Uplift models are now on their 4th generation and continuously evolving. Faster

model refreshes (from weeks to days) mean they can run more campaigns efficiently.

🎯 **Precision Targeting**
Uplift modeling identifies customers who would only respond positively if treated, and avoids “sleeping dogs” (customers who might churn if disturbed).

🏦 **Broader Impact**
Success in Home Equity campaigns sparked expansion into direct deposits and other banking products, showing uplift modeling's flexibility.