AI in Marketing: The Value of Predictions

Predictions Have Economic Value

- This why companies are so eager to adopt machine learning technology
- However, each application is different.
  - Prediction quality varies.
  - Economic value varies.
- Important to evaluate the economic value of predictions for each application.

The Economic Value of Predictions

Basic Example: Targeted Advertising

Problem Settings 1

- The value of a new customer is \$5.00 (e.g., Customer Lifetime Value)
- Advertising impressions cost \$0.01
- 0.1% of ad impressions are clicked and lead to a purchase

1. **No Prediction**: What profit do we expect from showing the ad to a pool of 1000 random consumers?
2. **Perfect Prediction**: What profit do we expect from of showing the ad to only the consumers in the pool of 1000 who will click?

Expected Value Calculation

Basic Example: Targeted Advertising

1. No Prediction: Show the ad to all 1000 consumers

- Cost of impressions $= \$0.01 \times 1000 = \$10$
- Value of new customers $= \$5.00 \times 1000 \times 0.001 = \$5$
- Profit $= \$5 - \$10 = -\$5$

2. Perfect Prediction: Show the ad **only** to customers who click

- Cost of impressions $= 1000 \times 0.001 \times \$0.01 = \$0.01$
- Value of new customers $= \$5.00 \times 1000 \times 0.001 = \$5$
- Profit $= \$5 - \$0.01 = \$4.99$

Expected Value Calculation

1. Assign a \$ value v to each possible outcome o, represented as $v(o)$
2. Multiply each outcome by its probability, represented as $p(o)$
3. Sum each of the products:

$$\text{Expected Value} = p(o_1) \times v(o_1) + p(o_2) \times v(o_2) + p(o_3) \times v(o_3) \ldots$$

Where do the $p(o)$ and $v(o)$ values come from?

- $v(o)$ requires business knowledge of a given application.
- $p(o)$ can be estimated using the model and data.

The Economic Value of Predictions: Practice Problem

Targeted Advertising

Problem Settings 2

- The value of a new customer is $400.00 (e.g., Customer Lifetime Value)
- Advertising impressions cost $1.00
- 0.025% of ad impressions are clicked then lead to a purchase.

1. **No Prediction**: What profit do we expect from showing the ad to a pool of 1000 random consumers?
2. **Perfect Prediction**: What profit do we expect from of showing the ad to only the consumers in the pool of 1000 who will click?

Expected Value Calculation: Practice Problem

Targeted Advertising

1. No Prediction: Show the ad to all 1000 consumers

- Cost of impressions $= \$1 \times 1000 = \$1000$
- Value of new customers $= \$400 \times 1000 \times 0.00025 = \$100$
- Profit $= \$100 - \$1000 = -\$900$

2. Perfect Prediction: Show the ad **only** to customers who click

- Cost of impressions $= 1000 \times 0.00025 \times \$1 = \$0.25$
- Value of new customers $= \$400 \times 1000 \times 0.00025 = \$100$
- Profit $= \$100 - \$5 = \$99.75$

Real-World Predictions Are **Not** Perfect

Instead, your model will get some predictions wrong

Prediction quality can be represented with a **Confusion Matrix**
This table is for binary outcomes: Predictions and actual outcomes are either 0 or 1.
Example:

	Predicted 0	Predicted 1	Total
Actual 0	598	79	677
Actual 1	103	220	323
Total	701	299	1000

Now how do we calculate profits?

Real-World Predictions Are **Not** Perfect

Calculate probabilities for each outcome using the Confusion Matrix

	Predicted 0	Predicted 1	Total
Actual 0	$N_1 = 598$	$N_3 = 79$	677
Actual 1	$N_2 = 103$	$N_4 = 220$	323
Total	701	299	1000

$\Downarrow$ convert to probabilities

	Predicted 0	Predicted 1
Actual 0	$p(o_1) = 0.598$	$p(o_3) = 0.079$
Actual 1	$p(o_2) = 0.103$	$p(o_4) = 0.220$

$$p(o_i) = \frac{N_i}{N_1 + N_2 + N_3 + N_4}$$

Not All Prediction Errors are Equal: Cancer Example

Consider two different errors a cancer screening test might make:

1. Test indicates cancer, but the person does **not** have it

- "False positive" result
- Causes the patient a great deal of stress
- The patient must return for additional testing
- Eventually, the additional testing reveals no cancer

2. Test indicates **no** cancer, but the person **does** have it

- "False negative" result
- The patient does not return for additional testing
- Missing early detection threatens the patient's life

Which outcome is worse? Should we weight these errors equally?

Not All Prediction Errors are Equal: Marketing Example

Consider two different errors a cancer screening test might make:

1. Model predicts the consumer will purchase, but they don't

- "False positive" result
- Have to pay for the impression.
- The customer generates no revenue.
- Negative profit value for this outcome.

2. Model predicts the consumer will not purchase, but they would have

- "False negative" result
- Don't have to pay for the impression
- Don't receive any revenue.
- Zero effect on profits.

Which outcome is worse? Should we weight these errors equally?

The Cost-Benefit Matrix

	Predicted 0	Predicted 1
Actual 0	$v(o_1)$	$v(o_3)$
Actual 1	$v(o_2)$	$v(o_4)$

- $v(o_1)$ is the value the company gets from correctly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the customer won't click, then don't serve him the ad.
- $v(o_2)$ is the value the company gets from incorrectly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the company predicts the customer will not click it does not serve an impression.
- $v(o_3)$ is the value the company gets when they incorrectly predict that a consumer will take an action.
  - *Often, this is a negative* number. If the company serves an ad and the customer doesn't respond, the company still has to pay fo the ad.
- $v(o_4)$ is the value the company gets when they correctly predict the consumer will taken an action.
  - *Often, this is a positive number*. These are the occasions in which the company makes money off of customers.

The Cost-Benefit Matrix

	Predicted 0	Predicted 1
Actual 0	$v(o_1)$	$v(o_3)$
Actual 1	$v(o_2)$	$v(o_4)$

- $v(o_1)$ is the value the company gets from correctly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the customer won't click, then don't serve him the ad.
- $v(o_2)$ is the value the company gets from incorrectly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the company predicts the customer will not click it does not serve an impression.
- $v(o_3)$ is the value the company gets when they incorrectly predict that a consumer will take an action.
  - *Often, this is a negative* number. If the company serves an ad and the customer doesn't respond, the company still has to pay fo the ad.
- $v(o_4)$ is the value the company gets when they correctly predict the consumer will taken an action.
  - *Often, this is a positive number*. These are the occasions in which the company makes money off of customers.

The Cost-Benefit Matrix

	Predicted 0	Predicted 1
Actual 0	$v(o_1)$	$v(o_3)$
Actual 1	$v(o_2)$	$v(o_4)$

- $v(o_1)$ is the value the company gets from correctly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the customer won't click, then don't serve him the ad.
- $v(o_2)$ is the value the company gets from incorrectly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the company predicts the customer will not click it does not serve an impression.
- $v(o_3)$ is the value the company gets when they incorrectly predict that a consumer will take an action.
  - *Often, this is a negative* number. If the company serves an ad and the customer doesn't respond, the company still has to pay fo the ad.
- $v(o_4)$ is the value the company gets when they correctly predict the consumer will taken an action.
  - *Often, this is a positive number*. These are the occasions in which the company makes money off of customers.

The Cost-Benefit Matrix

	Predicted 0	Predicted 1
Actual 0	$v(o_1)$	$v(o_3)$
Actual 1	$v(o_2)$	$v(o_4)$

- $v(o_1)$ is the value the company gets from correctly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the customer won't click, then don't serve him the ad.
- $v(o_2)$ is the value the company gets from incorrectly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the company predicts the customer will not click it does not serve an impression.
- $v(o_3)$ is the value the company gets when they incorrectly predict that a consumer will take an action.
  - *Often, this is a negative* number. If the company serves an ad and the customer doesn't respond, the company still has to pay fo the ad.
- $v(o_4)$ is the value the company gets when they correctly predict the consumer will taken an action.
  - *Often, this is a positive number*. These are the occasions in which the company makes money off of customers.

The Cost-Benefit Matrix

	Predicted 0	Predicted 1
Actual 0	$v(o_1)$	$v(o_3)$
Actual 1	$v(o_2)$	$v(o_4)$

- $v(o_1)$ is the value the company gets from correctly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the customer won't click, then don't serve him the ad.
- $v(o_2)$ is the value the company gets from incorrectly predicting that the consumer will not take an action.
  - *Often takes value zero*: E.g., if the company predicts the customer will not click it does not serve an impression.
- $v(o_3)$ is the value the company gets when they incorrectly predict that a consumer will take an action.
  - *Often, this is a negative* number. If the company serves an ad and the customer doesn't respond, the company still has to pay fo the ad.
- $v(o_4)$ is the value the company gets when they correctly predict the consumer will taken an action.
  - *Often, this is a positive number*. These are the occasions in which the company makes money off of customers.

Baseline for the Cost-Benefit Matrix

- The values in the cost-benefit matrix should be defined relative to a baseline.
- A convenient baseline is: "What if the company does nothing?"
  - e.g., instead of targeted advertising, what if the company does no advertising
  - e.g., instead of managing churn, what if the company does not manage churn
- With this baseline, the results are easy to interpret:
  - e.g, "How much better is this targeted advertising relative to not advertising?"
- Other baselines are possible, but may be harder to interpret.

Assign values to the Cost-Benefit Matrix

Assign values to the Cost-Benefit Matrix

- $v(o_i) = \$\text{Revenue}_i - \$\text{Cost}_i$
- Using Problem Settings 1, for example:

	Predicted 0	Predicted 1
Actual 0	$v(o_1) = \$0 - \$0 = \$0$	$v(o_3) = \$0 - \$0.01 = -\$0.01$
Actual 1	$v(o_2) = \$0 - \$0 = \$0$	$v(o_4) = \$5 - \$0.01 = \$4.99$

- Cost-Benefit matrix has the same rows and columns as the Confusion Matrix
- Values in each cell come from business knowledge of the application.

Calculate Expected Value

Confusion Matrix Probabilities

	Predicted 0	Predicted 1
Actual 0	$p(o_1) = 0.598$	$p(o_3) = 0.079$
Actual 1	$p(o_2) = 0.103$	$p(o_4) = 0.220$

Cost-Benefit Matrix (using Problem Settings 1 here)

	Predicted 0	Predicted 1
Actual 0	$v(o_1) = \$0 - \$0 = \$0$	$v(o_3) = \$0 - \$0.01 = -\$0.01$
Actual 1	$v(o_2) = \$0 - \$0 = \$0$	$v(o_4) = \$5 - \$0.01 = \$4.99$

Expected Value (EV) of the Model's Predictions

$$EV = p(o_1) \times v(o_1) + p(o_2) \times v(o_2) + p(o_3) \times v(o_3) + p(o_4) \times v(o_4)$$
$$EV = 0.598 \times \$0 + 0.103 \times \$0 + 0.079 \times (-\$0.01) + 0.220 \times \$4.99$$
$$EV = \$1.10 = \text{Average profit from a potential customer.}$$

Expected Value Calculations: Why are they important?

- These are helpful for determining whether a given marketing application is likely to benefit the company.
  - Positive numbers suggest its profitable
  - Negative numbers suggest it will hurt the company
- Expected value calculations help decide which model to use
  - It's hard to anticipate which predictive model will work best for a given application
  - Expected value provides a clear measure for which would work best for a given application.

The Cost-Benefit Matrix: Political Considerations

- The cost-benefit matrix can have a strong influence on which model is used.
- The values in the cost-benefit matrix should be agreed on before performing other analysis.
- If not, some functions may claim the results are not valid because they don't agree with the cost-benefit values.
- Prevent this by getting buy-in ahead of time.
- Cost-benefit parameters are less political: they don't have implications for how the company should proceed on their own.
  - Cost-benefit values only influence modeling decisions when combined with the complete analysis.

Assign values to a new problem: Customer Churn

	Predicted 0	Predicted 1
Actual 0	o_1	o_3
Actual 1	o_2	o_4

Meaning: Will Maintenance Prevent Churn?

- "Respond to maintenance" includes not churning due to a coupon, gift, special discount or personalized email.
- o_1: Correctly predicted that the customer would **not** respond
- o_2: Incorrectly predicted that the customer would **not** respond
- o_3: Incorrectly predicted that the customer **would** respond
- o_4: Correctly produced that the customer **would** respond

How should we value each outcome?

- What are the key costs and revenues associated with each outcome?

Assign values to a new problem: Customer Churn

	Predicted 0	Predicted 1
Actual 0	$v(o_1)$	$v(o_3)$
Actual 1	$v(o_2)$	$v(o_4)$

Cost-benefits relative to doing nothing to prevent churn:

- $v(o_1) = \$0$. Nothing spent, no change to the customers status.
- $v(o_2) = \$0$. Nothing spent. Customer would have churned anyway.
- $v(o_3) = -\$\text{Maintenance cost}$. Spent maintenance cost, but the customer would not have churned.
- $v(o_4) = \$\text{CLV} \times \text{Maintenance response share} - \$\text{Maintenance cost}$.
  - Spent maintenance cost and retained the share of customers who respond to maintenance.

Practice Problem

Basic Example: Preventing Churn

Problem Settings 3

- The value of a new customer is \$10 (e.g., Customer Lifetime Value)
- Personalized emails to prevent churn cost \$1 each.

Confusion Matrix:

	Predicted 0	Predicted 1
Actual 0	8000	2000
Actual 1	3000	1000

What is the expected value using these predictions to prevent churn?

Calculate Expected Value

Confusion Matrix Probabilities

	Predicted 0	Predicted 1
Actual 0	$p(o_1) = 0.57$	$p(o_3) = 0.14$
Actual 1	$p(o_2) = 0.21$	$p(o_4) = 0.13$

Cost-Benefit Matrix (using Problem Settings 3 here)

Cost-Benefit matrix relative to doing nothing to prevent churn

	Predicted 0	Predicted 1
Actual 0	$v(o_1) = \$0$	$v(o_3) = -\$1$
Actual 1	$v(o_2) = \$0$	$v(o_4) = \$10 - \$1 = \$9$

Expected Value (EV) of the Model's Predictions

$$EV = p(o_1) \times v(o_1) + p(o_2) \times v(o_2) + p(o_3) \times v(o_3) + p(o_4) \times v(o_4)$$
$$EV = 0.57 \times \$0 + 0.21 \times \$0 + 0.14 \times (-\$1) + 0.13 \times \$9$$
$EV = \$1.03 =$ Average profit per customer from churn targeting.

Common Names for the Confusion Matrix

Often, these names are used for each of the outcomes we studied:

	Predicted 0	Predicted 1
Actual 0	True Negative (TN)	False Positive (FP)
Actual 1	False Negative (FN)	True Positive (TP)

Polling questions

Access the poll here:

www.pollEv.com/mthomas

- Log in with your NUS ID
- This will allow me to give credit for your participation.

Practice Questions

Predictions have economic value because:

- Ans: They help businesses to spend their marketing dollars more efficiently.
- They inform about the business value of different outcomes in the confusion matrix
- They inform us how to assign different values to different types of errors.
- They avoid politically-driven decisions.

Practice Questions

Assigning values to the cost-benefit matrix requires an understanding of:

- Ans: How the predictions will be applied.
- The company's balance sheet.
- The company's cash flow statement.
- The company's cost of borrowing money.
- Which model will be used to make the predictions.

Practice Questions

Which of the following is likely to be the most politically charged:

- Ans: The choice of analytical approaches.
- The cost-benefit matrix
- The confusion matrix.
- The customer lifetime value

Practice Questions

Expected value calculations will help us to:

- Ans: Estimate which analytical approach is most effective.
- Estimate the cost-benefit matrix
- Estimate the confusion matrix.
- Estimate the customer lifetime value.

Practice Questions

If we are trying to predict which customers to target with advertising, which cell in this cost-benefit matrix is likely to have a positive value:

	Predicted 0	Predicted 1
Actual 0	Cell 1	Cell 3
Actual 1	Cell 2	Cell 4

- Ans: Cell 4
- Cell 1
- Cell 2
- Cell 3
- None of these. All values are ≤ 0.

Practice Questions

Confusion Matrix:

	Predicted 0	Predicted 1
Actual 0	800	100
Actual 1	30	20

Cost-benefit Matrix:

	Predicted 0	Predicted 1
Actual 0	0	-$0.10
Actual 1	0	$5.00

What is the expected value of these predictions?

$$\text{EV} = 0.80 \times 0 + 0.03 \times 0 + 0.10 \times (-\$0.10) + 0.02 \times \$5 = \$0.09$$

Practice Questions

Imagine you are trying to implement a targeted advertising strategy. To determine the value of Cell 3 in the cost-benefit matrix, which of the following would be most helpful:

	Predicted 0	Predicted 1
Actual 0	Cell 1	Cell 3
Actual 1	Cell 2	Cell 4

- Ans: The cost of advertising impressions
- The customer lifetime value
- The churn rate.
- The share of impressions that receive a click.
- The share of impressions that lead to a purchase.
- None of these