

The background of the slide is a dark, grainy image of an NHL arena. A player in a blue jersey is visible in the center, holding a hockey stick. The arena floor and some spectators in the background are faintly visible.

NHL PLAYER VALUATION AND TRADE RECOMMENDATION SYSTEM

SANGMUN KIM, YUYANG CHEN, HAICHEN SUN, MICHAEL KUBY

CHALLENGE: PREDICT THE SALARY OF A PLAYER BASED ON PERFORMANCE METRICS

RECALL: $Y = Xw$

X : PERFORMANCE METRICS FOR 2007 - PRESENT

Y : SALARY FOR 2007 – PRESENT

w : TO BE LEARNED

```
[>>> df.head()
  Position Team      Player      TOI/GP  GP  ... Draft Team Draft Round Round Pick Overall Draft Position  Year
0         R  DET  Aaron Downey  4.5880952380952  56  ...      -      -      -      -      -      -      2007
1         D  NYI  Aaron Johnson  13.8616666666667  30  ...    CBJ      3     22      85      2007
2         D  VAN  Aaron Miller  17.336549707602  57  ...    NYR      5      4     88      2007
3         D  CBJ   Aaron Rome  18.178431372549  17  ...    L.A      4      8    104      2007
4         L  MIN  Aaron Voros  9.1745454545455  55  ...    N.J      8      2    229      2007

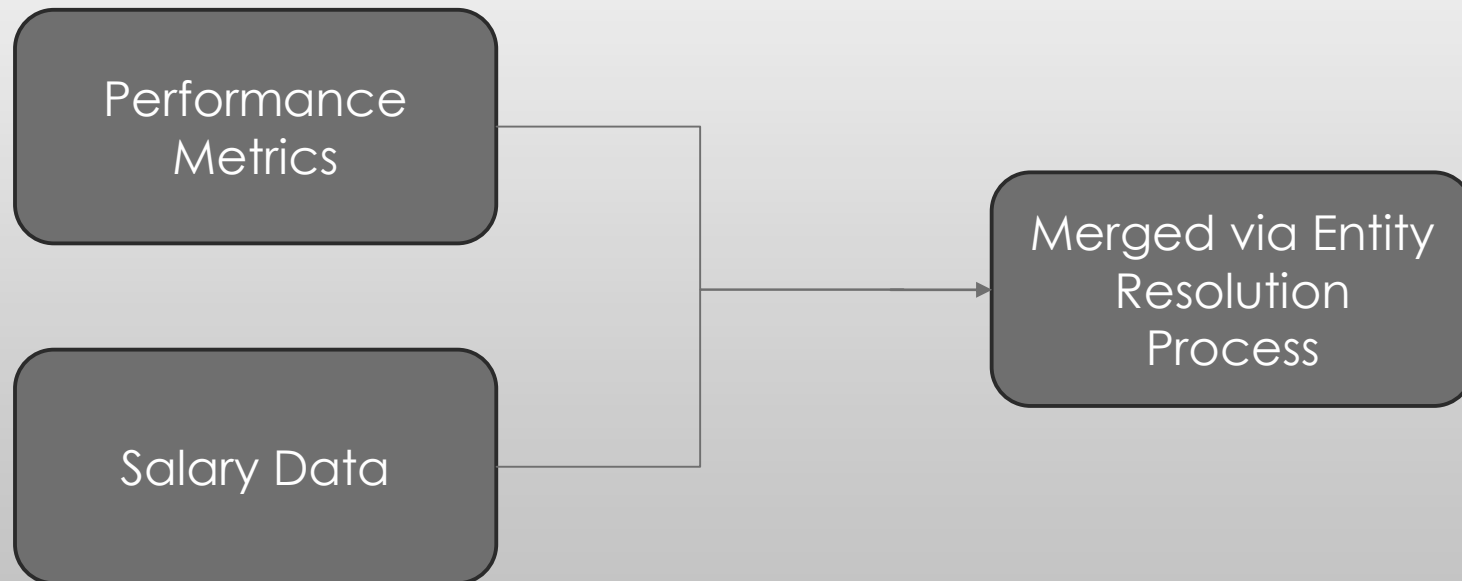
[5 rows x 97 columns]
```

WORKFLOW STEP 1: PROCURE THE DATA

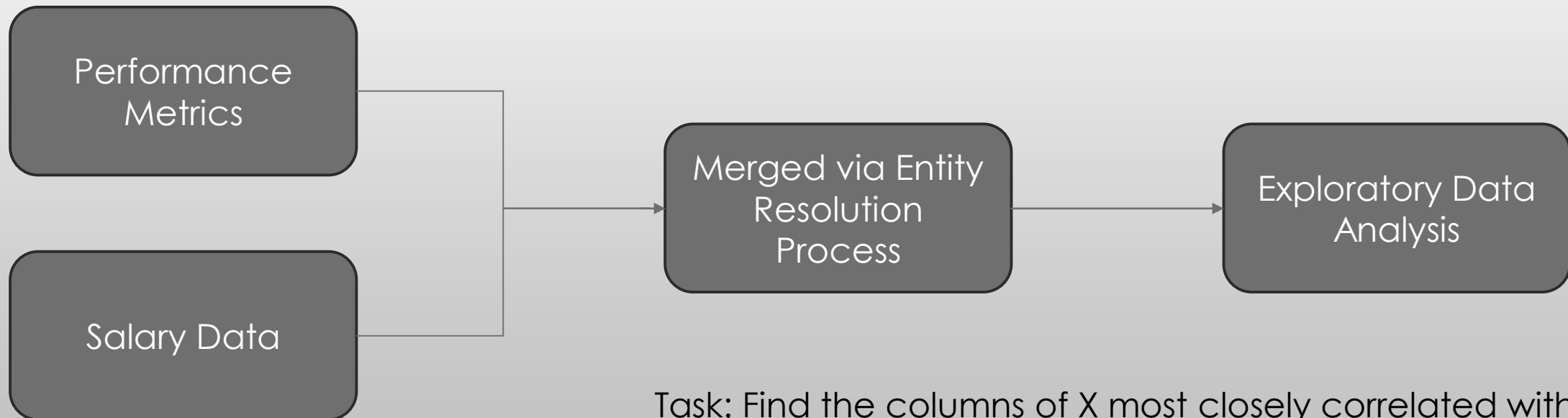
X (PERFORMANCE METRICS): SCRAPED FROM WWW.NATURALSTATTRICK.COM

Y (SALARY DATA): SCRAPED FROM WWW.CAPFRIENDLY.COM

WORKFLOW STEP 2: ENTITY RESOLUTION



WORKFLOW STEP 3: EXPLORATORY DATA ANALYSIS



Task: Find the columns of X most closely correlated with Y.
Tools: Correlation Analysis, Matplotlib/Seaborn
Other: Principal Component Analysis

WORKFLOW STEP 4: TRAIN A REGRESSION MODEL



Decisions and Challenges:

- Multiple models (one for each position)?
- Format for our Y's (how to adjust for inflation)?
- Tools (AutoML)?

WORKFLOW STEP 5: PREDICT

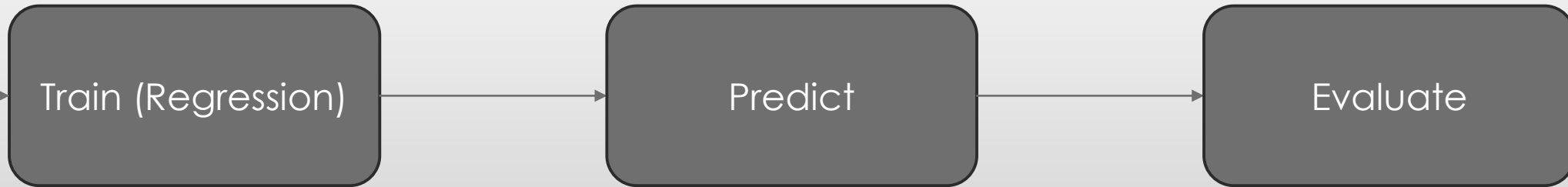


Understanding the predictions:

“Predictions are what the model thinks the player is *worth*, given their *performance*.”

A measure of **value**.

WORKFLOW STEP 6: PLAYER EVALUATION



Consider: Every player **already has a salary**.

Recall: Prediction = value

Player Evaluation:

If Value – Salary > 0:
Player is **overperforming**

If Value – Salary < 0:
Player is **underperforming**

WORKFLOW STEP 7A: RECOMMEND



Recommend: Given a player that we would like to trade, what players should we **target**?

Player to trade: **underperformer**

Player to target: **overperformer**

Constraint: Should have a comparable **salary**.

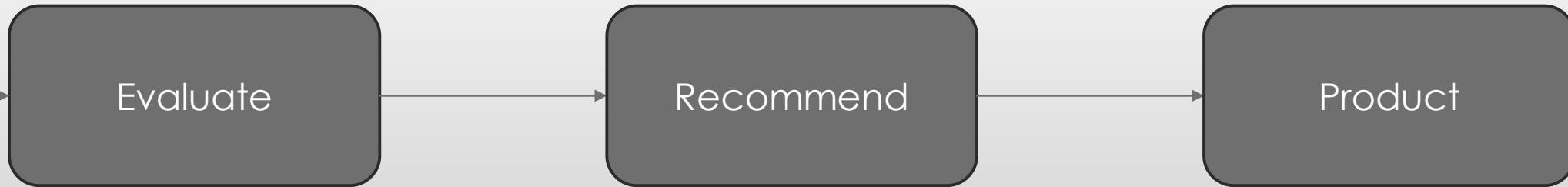
WORKFLOW STEP 7B: RECOMMEND



Time Permitting: Include sentiment analysis by analysis /r/{hockey, teamA, teamB, ...}

Players to target: **overperformer + sentiment analysis**

WORKFLOW STEP 8: PRODUCT



Product: A web-based interface

Features:

1. Select Team → Select Player → View valuation and relevant metrics.
2. Produce a list of players identified as possible trade targets.

A person wearing a blue jacket, orange backpack, and dark pants is kneeling on a snowy slope. They are holding a long, thin pole or stick that extends diagonally across the frame. The background is a dark, textured surface, possibly a wall or a backdrop, with a subtle grid pattern. The overall scene is dimly lit, with the person's clothing and the pole providing some contrast against the dark background.

THANK YOU!