



Predicting Tommy John Surgery

John Streaman





Business Understanding

Collect data on MLB pitchers to anticipate risk for TJ surgery before signing, & mitigate practices that may result in needing surgery.





Intro: Who is Tommy John?

First player to receive surgery on UCL (1974)

**Replace with tendon from persons forearm/
hamstrings**

- **Before this, considered career ending injury (dead arm)**
- **Had a successful career (retired 1989)**





Protect Your Investment

Signed record deal - 10 Year, \$700 Million

Potential Fees:

- Luxury tax
- (International signing fees)
- Insurance
- Ticket sales, merch, etc.

First TJ Surgery - 2018

September 2023

- Shutdown for injury
- MRI - torn UCL
- Underwent 2nd TJ Surgery



Shohei Ohtani injury update: Angels superstar shut down for season, likely getting elbow surgery soon

Ohtani's on-field Angels career could be finished, but he will be around the team for the rest of the season



By [Mike Axia](#) Sep 18, 2023 at 5:13 pm ET • 2 min read





Data

General pitching statistics (1972-2023) from baseball-reference.com

Running Google doc of baseball players who received TJ surgery

- Advanced pitching metrics gathered from [pybaseball](#)
- Collected, cleaned, and merged to create a dataset



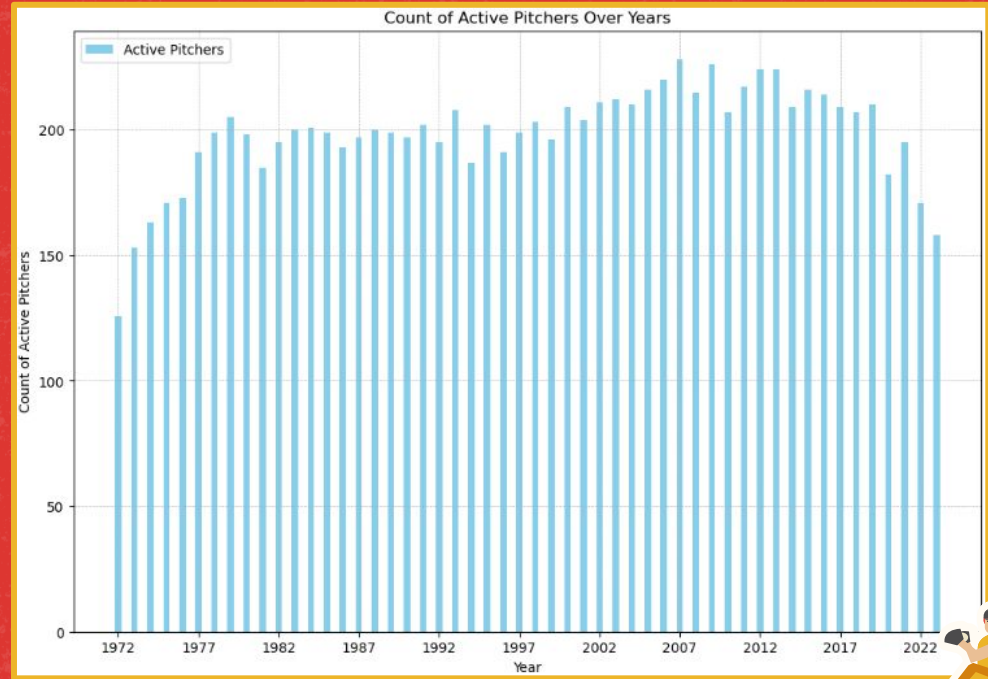
Historic Data

1972 - 2023

Starting Pitchers only

Minimum 320 career innings thrown

- Early 2000's
- Dramatic rise in number of pitchers

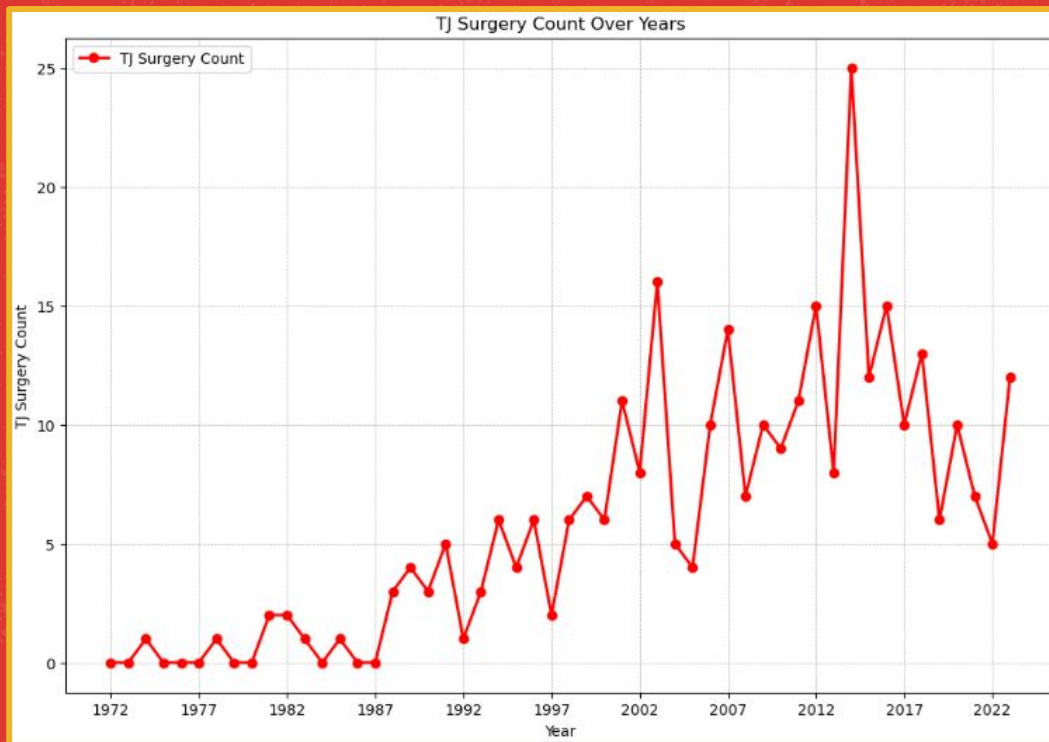


Historic Data cont.

TJ Surgery Count

Dramatic spikes:

- 2003
 - 'Steroid Era'
- 2014
 - Increase in younger generation play, higher velocity



Advanced Pitching Metrics

Seasonal data - limited

- 2008-2023
- Starting Pitchers, min. 320 innings (career)

Features - 129

Target - 'Surgery'

- Binary (1.0 - yes, 0.0 - no)
- Count imbalanced
 - 1.0 - 916
 - 0.0 - 2772

★ Grouped features for each pitcher by season

- Age, season, Throws (L/R)
- Total number thrown for each pitch type
 - fastball, curveball, slider, etc.
 - Avg. speed, velocity, & release position



Pitch Types

Fastball

- Fastball (FF)
- Sinker (SI)
- Cutter (FC)

Breaking ball

- Curveball (CU)
- Knuckle-curve (KC)
- Slow curve (CS)

Other

- Eephus (EP)
- Absent (AB)
- Intentional Walk (IN)

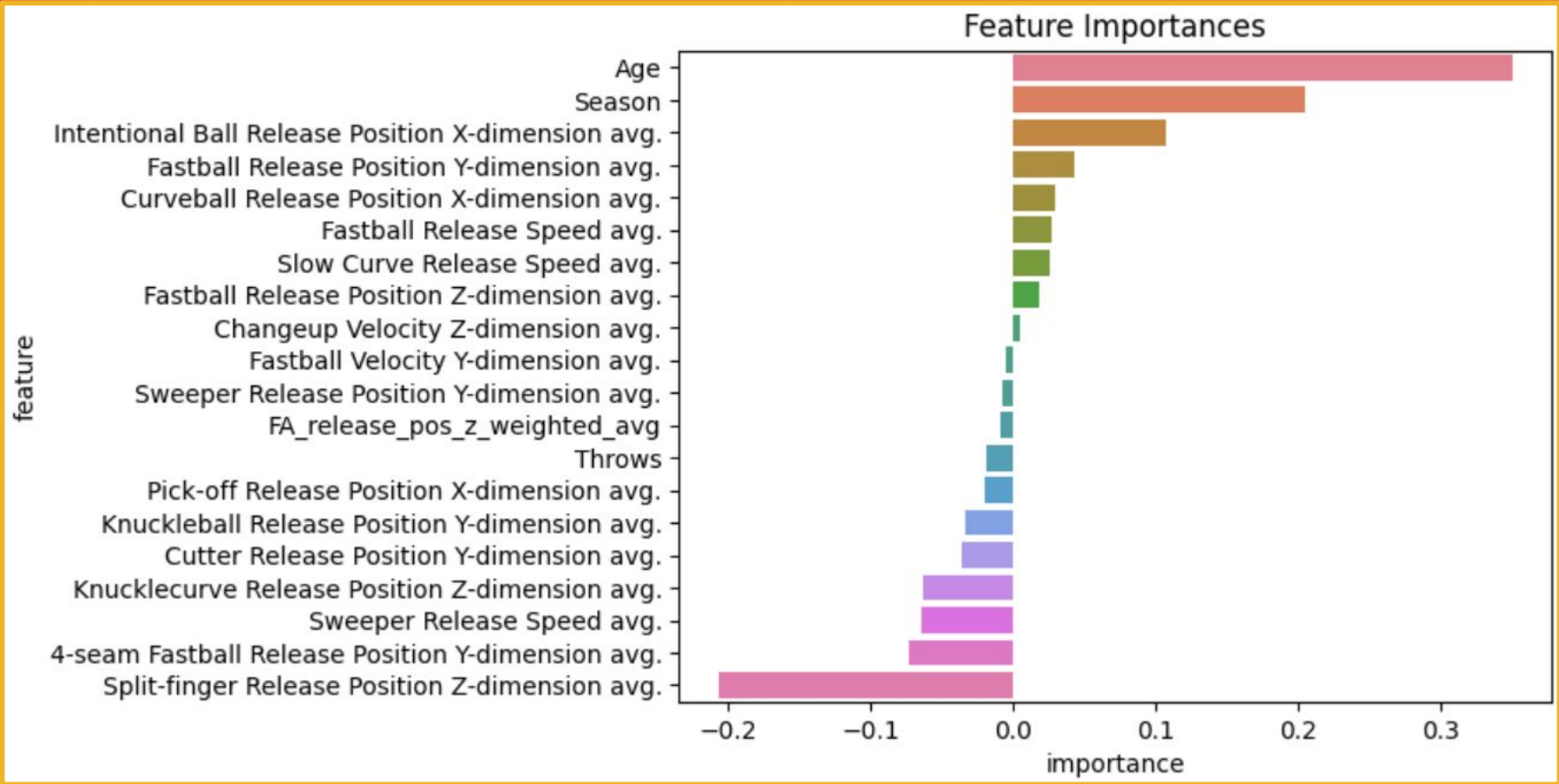
Off-speed

- Change-up (CH)
- Split-finger (FS)
- Forkball (FO)
- Screwball (SC)

Slider

- Slider (SL)
- Sweeper (ST)
- Slurve (SV)
- Knuckle (KN)

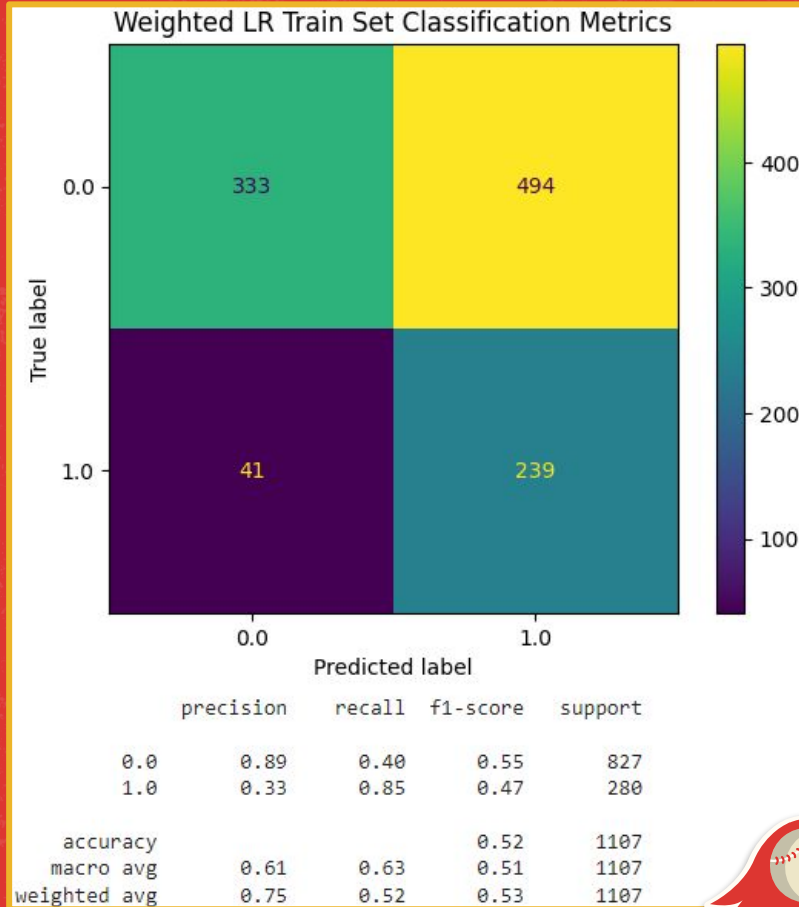




Best Model

Logistic Regression Model

Adjusted class weights



Next Steps

Collect more data!

- Advanced pitching data prior 2008 (velocity, rotation, etc.)
- Video analysis - pitching mechanics
- Biological data (height, weight, diet, genetics, etc.)
- International leagues, minors, spring training, bullpen, college, etc.
- Other factors
 - Days rest, temperature, humidity, altitude, etc.





Questions?

