Video Segment 3: Quality Control: Data Integrity, Bias and Statistical Models

"The data you wish you had is never coming." -Dr. Eric Bradlow, Wharton Professor

NOTE-TAKING GUIDE

Professor Eric Bradlow notes two issues that academics and others are focused on to ensure the data integrity of sports. Consult your vocabulary list for a definition of **data integrity**, and then take note of Dr. Bradlow's two areas of interest in the video segment:

1.

2.

Relating to data privacy and the use of high-integrity data, what is the relevance of:

The KYC Process:

Sportsradar and Bet Genius:

Decision-making at the Philadelphia Eagles:

WHARTON
UNIVERSITY of PENNSYLVANIA
GLOBAL YOUTH PROGRAM

Explore Business: More Than a Game | 1
©2023 The Wharton School, The University of Pennsylvania
globalyouth.wharton.upenn.edu

Professor Bradlow says, "Every [machine learning] model has observable factors and unobservable factors." How do the following people and organizations address these factors?

Researchers at Wharton:

Risk and Trading Organization at FanDuel:

Reflection: This video segment in part talks about online data privacy. Information about us is generated as we go through the world and as we interact with government systems and private organizations – for example, anytime we buy something or interact with a business. Do you regularly consider your personal data and privacy when you are on social media or using a financial app? In what ways do you try to control your personal information? The world of Web3 may well address data privacy concerns by allowing us to have more control over the information we share online. How would you hope your information becomes more secure?

THE LANGUAGE OF BUSINESS
(Linked terms are defined in the Wharton Global Youth Video Glossary)

CEO or Chief Executive Officer

Company

Computer Science

Consumer

Data

Data Mining

Empirical

Industry

Marketing

MBA or Master of Business Administration

Public Policy

Regulated Market

Regulation

Statistician

Academic Research: Research conducted by faculty at a university like Penn that focuses on creating new ideas and perspectives. Academic researchers read all the existing studies and papers examining a specific topic or theory and then see where they can add value with new experimentation and insight.

Algorithm: A procedure used for solving a problem or performing a computation. Each algorithm is programmed as a step-by-step process for performing calculations and data processing.

Amazon: The world's largest online retailer.

Amy Howe: Chief Executive Officer (CEO) of FanDuel, who earned her Master of Business Administration (MBA) from the Wharton School in 1999.

Analytics: The systematic computational analysis of data or statistics. It is used for the discovery, interpretation, and communication of meaningful patterns in data.

Artificial Intelligence or AI: Intelligence demonstrated by machines (computers), rather than humans.

Bad Actor: A person or organization responsible for actions that are harmful, illegal, or morally wrong.

Betgenius: A provider of data-driven software to regulated sports betting and lottery operators.

Beyond Business: An ongoing conversation at the Wharton School of the University of Pennsylvania that explores the most complex and pressing issues impacting individuals and organizations across the world. Each year Wharton Global Youth creates Explore Business mini-sites for high school students around different Beyond Business themes.

Bias: Prejudice against or in favor of one thing, person or group compared with another, usually in a way that is unfair.

Bias in Machine Learning: When an algorithm produces results that are systematically prejudiced due to poor assumptions in the machine learning process – perhaps, for instance, the data is imperfect or biased.

Cybersecurity Threat: An act that threatens online security. Hackers, for example, might want to steal customer or company data or cause damage to a computer system.

Data Aggregation: Summarizing a large pool of data for high-level analysis.

Data Granularity: The measure of the level of detail in a set of data. Granular data is detailed data.

Data Integrity: The overall accuracy, completeness and consistency of data during its lifetime. Organizations work to maintain the integrity of the data they collect, so it can continue to be a valuable tool in supporting their decision-making.

Data Privacy: The ability of people to control their personal information, especially online. Organizations collect massive amounts of data on their customers and users and must work to keep that data secure.

Dataset: A collection of data.

Empirically Driven Decision Making: An empirical approach gathers observable data and uses evidence to study and interpret information – and ultimately arrive at results that support decision-making.

Eric Bradlow: A marketing professor at the Wharton School of the University of Pennsylvania and Vice Dean of Analytics at Wharton.

Erika James: Dean of the Wharton School with research interests in crisis leadership, workplace diversity and management strategy. Erika James hosts the annual Beyond Business series of business conversations around which Wharton Global Youth's Explore Business mini-sites are built.

FanDuel: An online gambling company founded in 2009 that offers sportsbook, daily fantasy sports, and more.

Flutter Entertainment: A global sports betting, gaming and entertainment company based in Ireland that owns and operates several brands, including FanDuel.

Howie Roseman: The executive vice president and general manager of the Philadelphia Eagles football team.

Jeffrey Lurie: Owner of the Philadelphia Eagles football team.

KYC: This stands for Know Your Customer and is a critical process for determining customer risk and whether a customer meets the requirements to use a company's services.

Machine Learning: Using and developing computer systems that learn and adapt by using algorithms and statistical models to analyze and understand patterns in data.

Parlay: A type of sports bet or wager in which two or more bets are linked together to create one bet that pays out more money.

Perturbing Data: A data security technique that adds "noise" to a database (a distortion) to protect sensitive information.

Philadelphia Eagles: A professional American football team based in Philadelphia, Pennsylvania.

Point Spread: The point margin between two teams in a game. A point spread bet is a wager on how much the favorite team will win by or the set number of points the underdog will not lose by for the wager to pay out.

Predictive Model: A statistical technique to predict future behavior. In this form of data-mining technology, you analyze historical and current data and generate a model to help predict future outcomes.

Risk and Trading Organization: The division of FanDuel that handles determining the odds and pricing for its bets.

Same-day Parlay: Combining two or more bets together in a single game with the hope of winning more money than if you had placed the bets individually.

Setting the Line: In gambling, this refers to creating the odds for a game, including determining the favorite and underdog team in a matchup. The oddsmaker relies on a complex set of mathematical models and computer algorithms to come up with the odds for a game.

Sophisticated Model: A data model is used to capture all the points of information necessary to conduct analysis; some models use sophisticated techniques and tools to discover deeper insights and make predictions.

Sportradar: A sports technology company that collects and analyzes sports data for bookmakers, sports federations and media.

Sports Analytics: The analysis of sports data, including information on player performance and business operations, to make predictions and support decision-making on and off the playing field.

Sports Betting: Predicting sports results, like who will win a football game, and placing a financial wager on the expected outcome in hopes of winning even more money.

Sports Betting Trader: The person who uses data, statistics and analysis to come up with the odds for a game or market that people want to bet on.

Sportsbook: An organization that takes bets on sporting events and pays out winnings.

Out-of-sample Predictive Accuracy

Statistical Model: A mathematical model that applies statistical analysis to a set of data.

Statistics: The practice or science of collecting, analyzing and interpreting numerical data in large quantities.

Tautology: When you say the same thing twice, just using different words.

Third-party Company: A company that provides a service for another company or for its customers.

The Wharton School: The Wharton School, also referred to as just Wharton, is the business school of the University of Pennsylvania, a private Ivy League university based in Philadelphia, Pennsylvania. Wharton is a world leader in business education and a hub for academic research in all areas of business.

EXPLORE MORE IN OUR ONLINE BUSINESS JOURNAL

The Power of Personal Data with Wharton's Kevin Werbach