AI in Statistics

MSI. Agustín Gómez Meléndez

UCR-UNED

November 1, 2023

Introduction

Figure: LINK VIDEO

Introduction

Prompt: AI in Statistics a new way of teaching

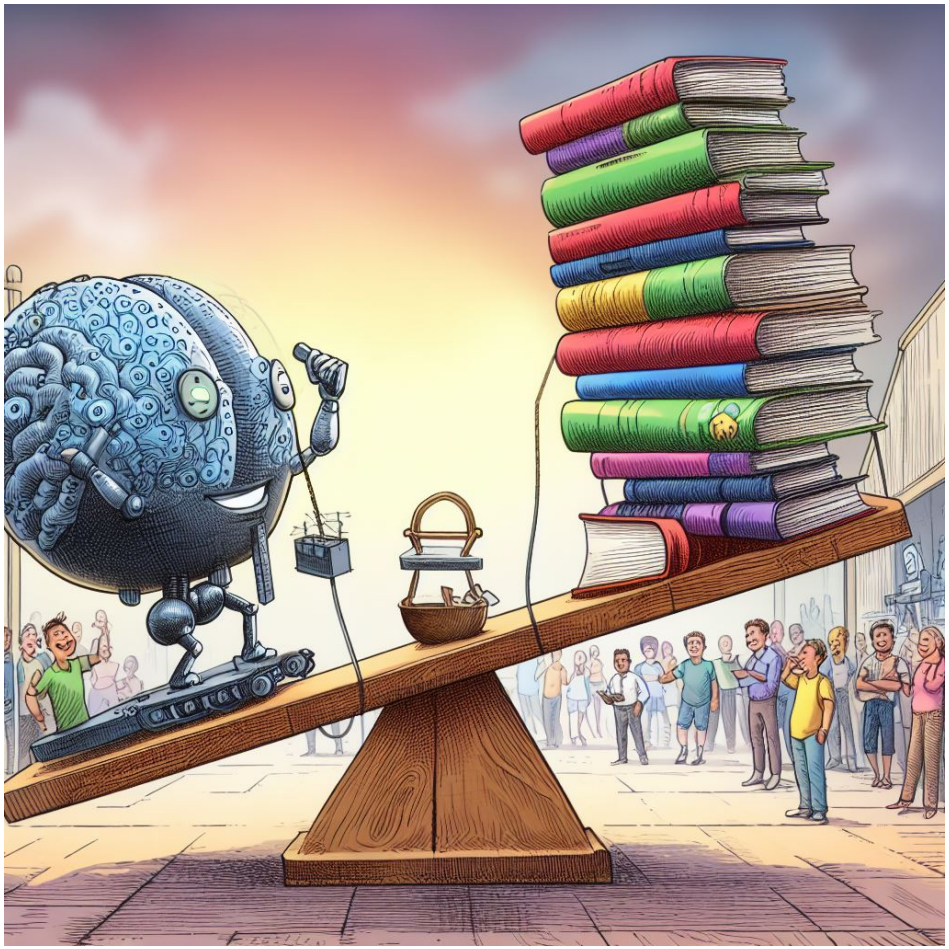


Figure: It happens before!

Introduction

What's the catch with GPT models?
"GPT" stands for "Generative Pre-trained Transformer." It is a type of artificial intelligence model developed by OpenAI.

- Generative: This refers to the model's ability to generate text or other content. GPT can create human-like text based on the patterns it has learned from the text it was trained on.

- Pre-trained: GPT is "pre-trained" on a vast amount of text data before it is fine-tuned for specific tasks. During pre-training, the model learns language structure, grammar, vocabulary, and some general knowledge from a diverse range of text sources.

- Transformer: The "Transformer" is a type of neural network architecture introduced in a paper titled "Attention is All You Need" by Vaswani et al. in 2017. Transformers have become the foundation for many state-of-the-art natural language processing models, including GPT. They use a mechanism called "attention" to process sequences of data, making them highly effective for tasks involving sequential data like text.

Introduction

Keywords: A new way of learn

1. Machine Learning	26. Overfitting	51. Facial Recognition	76. Fuzzy Logic
2. Neural Networks	27. Underfitting	52. Sentiment Analysis	77. Swarm Intelligence
3. Deep Learning	28. Cross-Validation	53. Text Generation	78. Evolutionary Algorithms
4. Natural Language Processing (NLP)	29. Training Data	54. Language Translation	79. Bayesian Networks
5. Computer Vision	30. Testing Data	55. Recommendation Systems	80. Knowledge Representation
6. Robotics	31. Validation Data	56. Anomaly Detection	81. Semantic Web
7. Reinforcement Learning	32. Model Evaluation	57. Computer Vision	82. Ontologies
8. Supervised Learning	33. Precision	58. Object Detection	83. Natural Language Understanding
9. Unsupervised Learning	34. Recall	59. Semantic Segmentation	84. Speech Synthesis
10. Semi-Supervised Learning	35. F1 Score	60. Edge Computing	85. Text-to-Speech
11. Algorithm	36. Accuracy	61. Internet of Things (IoT)	86. Chatbot
12. Data Mining	37. Loss Function	62. Big Data	87. Virtual Assistant
13. Pattern Recognition	38. Optimization	63. Cloud Computing	88. Machine Perception
14. Predictive Modeling	39. Backpropagation	64. Ethics in AI	89. Affective Computing
15. Artificial Neural Networks (ANN)	40. Hyperparameter Tuning	65. Bias in AI	90. Human-Computer Interaction
16. Convolutional Neural Networks (CNN)	41. Regularization	66. Explainable AI (XAI)	91. Augmented Reality
17. Recurrent Neural Networks (RNN)	42. Transfer Learning	67. Artificial General Intelligence (AGI)	92. Virtual Reality
18. Support Vector Machines (SVM)	43. Generative Adversarial Networks (GAN)	68. Artificial Superintelligence (ASI)	93. Simulated Reality
19. Decision Trees	44. Reinforcement Learning	69. Turing Test	94. Quantum Computing
20. Random Forests	45. Q-Learning	70. Machine Consciousness	95. Blockchain
21. Gradient Boosting	46. Deep Reinforcement Learning	71. Singularity	96. Smart Cities
22. K-Nearest Neighbors (KNN)	47. Chatbots	72. Data Science	97. Industry 4.0
23. Clustering	48. Autonomous Vehicles	73. Predictive Analytics	98. Health Informatics
24. Dimensionality Reduction	49. Speech Recognition	74. Cognitive Computing	99. Bioinformatics
25. Feature Engineering	50. Image Recognition	75. Expert Systems	100. Cyber-Physical Systems

Figure: Keywords

Introduction

Prompt: AI in Statistics a new way of teaching

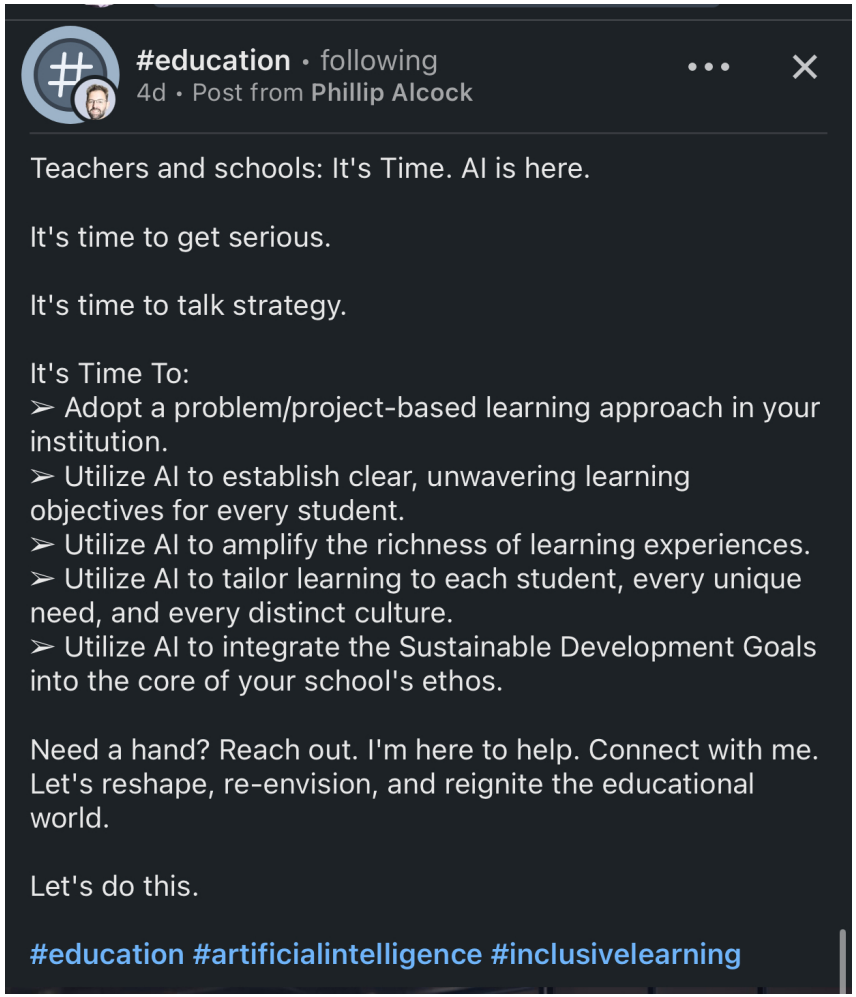


Figure: Taken from LinkedIn

Introduction

Prompt: AI in Statistics a new way of teaching

Figure: Created whit Dall-E

Model Architecture

Most competitive neural sequence transduction models have an encoder-decoder structure [5, 2, 29]. Here, the encoder maps an input sequence of symbol representations $(x_1, ..., x_n)$ to a sequence of continuous representations $\mathbf{z} = (z_1, ..., z_n)$. Given $\mathbf{z}$, the decoder then generates an output sequence $(y_1, ..., y_m)$ of symbols one element at a time. At each step the model is auto-regressive [9], consuming the previously generated symbols as additional input when generating the next.

The Transformer follows this overall architecture using stacked self-attention and point-wise, fully connected layers for both the encoder and decoder, shown in the left and right halves of Figure 1, respectively.

Figure: Attention Is All You Need

Model Architecture

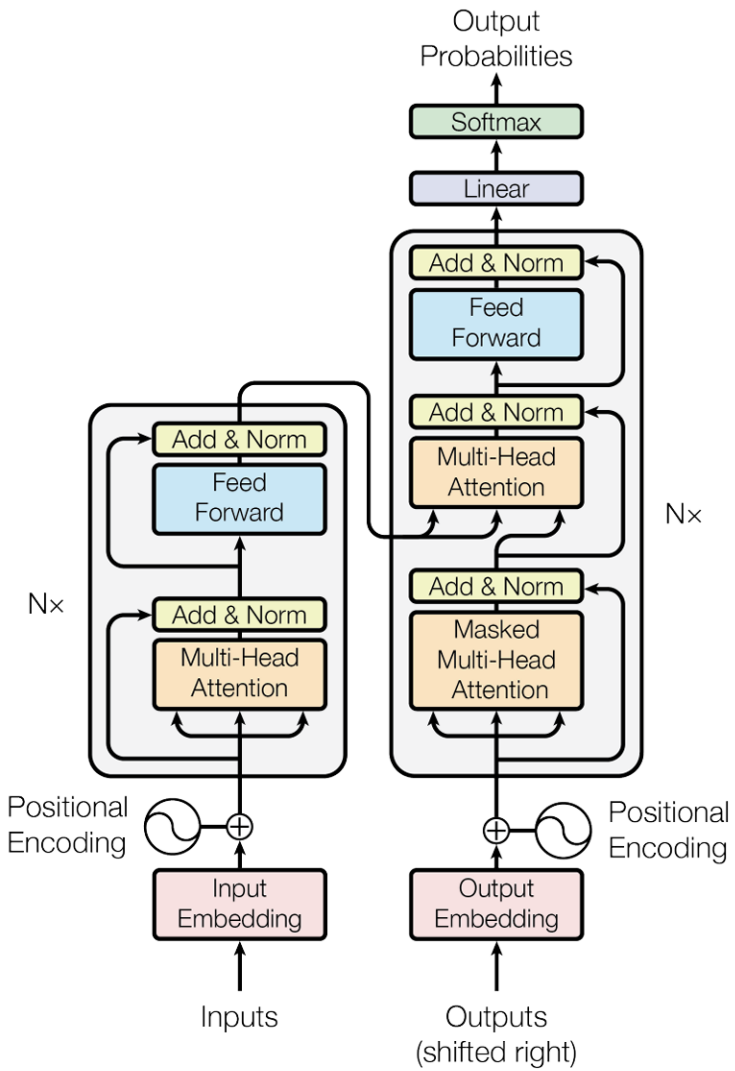


Figure 1: The Transformer - model architecture.

Introduction

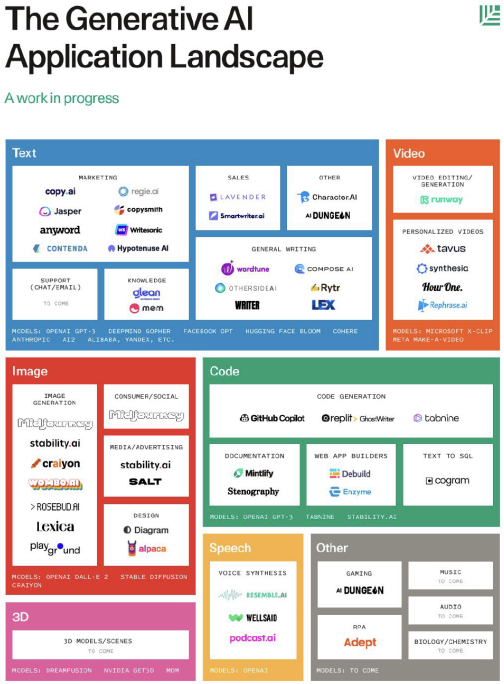
How Chatbots and Large Language Models Works

EXAMPLE 1 `https://youtu.be/X-AWdfSFCHQ`

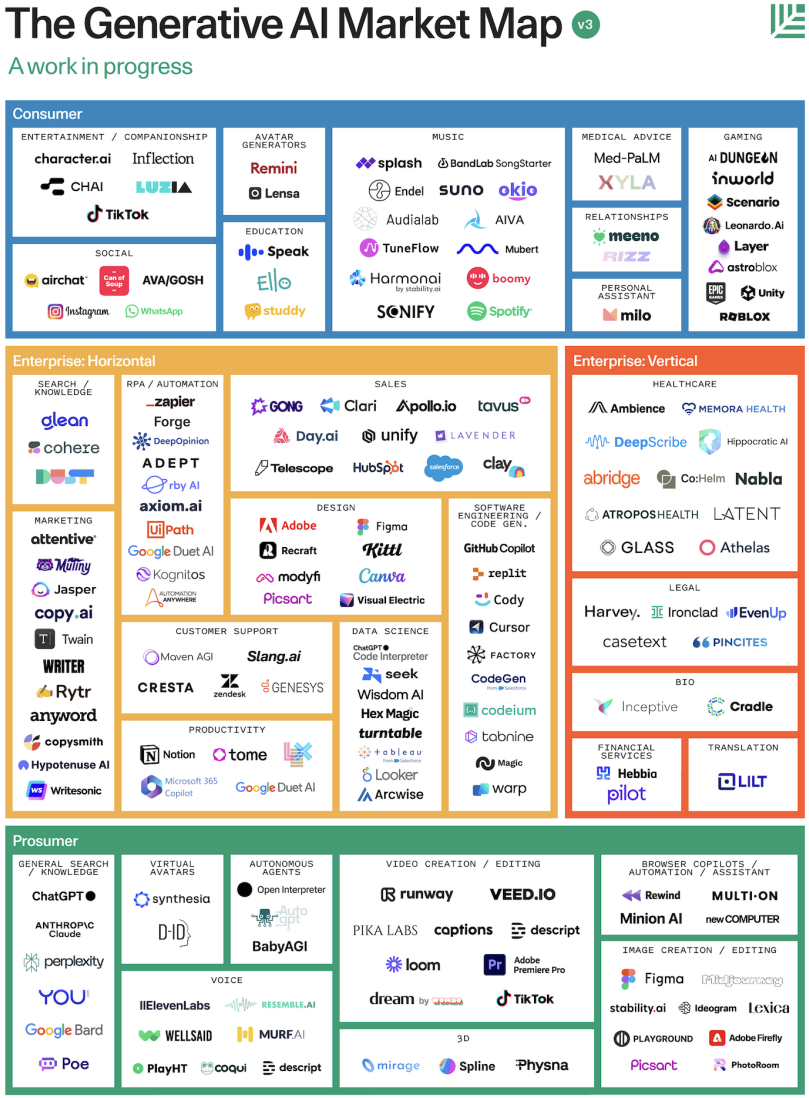
EXAMPLE 2 `https://www.youtube.com/watch?v=X994dDnmRmY`

Context

Article `https://www.sequoiacap.com/article/generative-ai-act-two/`



Marzo 2022



Infrastructure

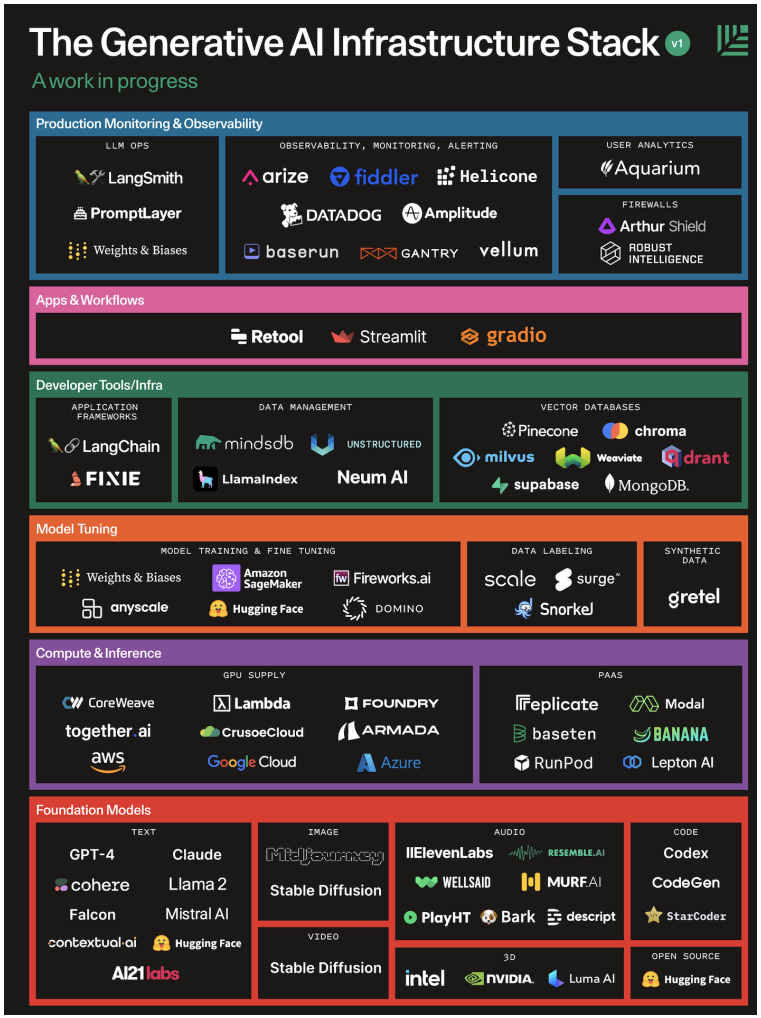


Figure: Source https://www.sequoiacap.com/

Definition and basic concepts of statistics

- Definition and basic concepts of statistics.
- Importance of statistics in decision-making
- Main Statistical Tools and Methods

Definition and basic concepts of statistics

Statistics is the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data. It serves as a tool to understand variability and offers methods to derive meaningful insights from numerical information.

The foundational concepts include differentiating between descriptive and inferential statistics, understanding the distinctions between populations and samples, and classifying variables as qualitative or quantitative

Importance of statistics in decision-making

Statistics plays a pivotal role in decision-making across various domains. It offers a structured way to represent data, helping individuals and organizations make informed decisions based on empirical evidence.

By understanding patterns, relationships, and trends within data, stakeholders can predict future outcomes, evaluate current strategies, and reduce uncertainties, ultimately leading to better, data-driven decisions

Main Statistical Tools and Methods

Statistical tools and methods encompass a wide range of techniques used to describe, analyze, and infer from data. Core tools include measures of central tendency (like mean, median, mode) and measures of dispersion (such as variance and standard deviation).

More advanced methods, such as hypothesis testing, regression analysis, and analysis of variance (ANOVA), help in making predictions or determining relationships between variables. The choice of tool or method often depends on the nature of the data and the specific research question.

Definition and basic concepts of artificial intelligence

(AI) refers to the simulation of human intelligence in machines, enabling them to perform tasks that typically require human-like thinking. These tasks include problem-solving, pattern recognition, decision-making, and language understanding, among others. Several foundational concepts underpin AI:

- **1. Algorithms:** At the heart of AI are algorithms, which are step-by-step procedures or formulas used to solve problems.

Definition and basic concepts of artificial intelligence

- **2. Machine Learning (ML):** A subset of AI, ML allows computers to learn from data without being explicitly programmed. For instance, given enough images of cats, a machine learning algorithm can identify a cat in a new image.

- **3. Neural Networks:** These are computing systems inspired by the structure of the human brain. They consist of interconnected nodes (analogous to neurons) and are foundational to deep learning, a powerful subset of ML.

Definition and basic concepts of artificial intelligence

- **4. Natural Language Processing (NLP)**: This involves enabling machines to understand and generate human language. It's the reason virtual assistants like Siri or Alexa can understand and respond to user commands.

- **5. Robotics**: A field of AI where machines can perform tasks in the real world. Examples include self-driving cars or robots in manufacturing.

Definition and basic concepts of artificial intelligence

- **6. Expert Systems:** These are computer systems that mimic the decision-making ability of a human expert in specific domains, using if-then rules and logic.

- **7. Knowledge Representation:** Storing information, facts, and rules that machines can use to understand the world, reason about it, and make decisions.

Advantages and benefits of artificial intelligence in education

- **Personalized Learning:** AI can analyze students' learning patterns and adapt content accordingly, offering a tailored learning experience. For example, if a student struggles with a particular math concept, AI-driven platforms can provide additional resources or exercises specifically in that area.

- **Efficiency and Automation:** Tasks such as grading assignments or tests can be automated using AI, allowing educators to spend more time on instruction and interaction with students. AI can also streamline administrative tasks for educational institutions.

Advantages and benefits of artificial intelligence in education

- **Data-Driven Insights:** AI can analyze vast amounts of data from students' interactions with digital platforms, providing insights into their learning behaviors, strengths, and areas of improvement. This can help educators refine their teaching strategies to better cater to their students' needs.

- **Enhancing Engagement:** AI-driven educational games and simulations can make learning more engaging and interactive. These tools can adjust in real-time to a student's actions, ensuring optimal challenge and engagement levels.

Advantages and benefits of artificial intelligence in education

- **Global Learning Environment:** AI-powered language translation tools can break down language barriers, enabling students from different parts of the world to access content in their native languages and facilitating global collaborative learning experiences.

- **Support for Diverse Needs:** AI can offer additional support for students with disabilities or learning differences. For instance, speech recognition can help students with writing difficulties, while personalized content can assist those with cognitive challenges.

Practical applications of artificial intelligence in the teaching of statistics

- **Adaptive Learning Platforms:** AI-driven educational software can adjust the learning material in real-time based on a student's progress and understanding. For example, if a student struggles with the concept of standard deviation, the platform can offer additional exercises or resources in that specific area until the student gains proficiency.

- **Data Analysis Tools:** AI can assist students in understanding large datasets by highlighting patterns, correlations, and other statistical phenomena, making data analysis more intuitive and less time-consuming.

Practical applications of artificial intelligence in the teaching of statistics

- **Simulations and Modelling:** AI can power simulations that allow students to visualize complex statistical concepts, such as distributions or hypothesis testing, by interacting with virtual models.

- **Automated Feedback:** AI-driven systems can provide instant feedback on statistical assignments, pointing out errors in calculations or interpretations and suggesting areas for improvement.

Practical applications of artificial intelligence in the teaching of statistics

- **Tutoring Systems:** AI can act as a personal tutor, providing students with answers to questions, guiding them through complex problems, and offering hints or solutions when they are stuck.

- **Real-World Data Collection:** Using AI, students can gather real-world data through sensors, social media scraping, or other methods for hands-on statistical analysis.

Practical applications of artificial intelligence in the teaching of statistics

- **Interactive Data Visualization:** AI tools can generate dynamic graphs and charts, allowing students to interact with data visualizations and gain deeper insights into the underlying statistics.
- **Statistical Language Processing:** For students working on textual data analysis, AI can help in parsing large volumes of text, identifying patterns, and extracting statistically relevant information.
- **Gamified Learning:** AI can drive games designed to teach statistical concepts. These games can adjust their difficulty based on a student's progress, ensuring an optimal learning curve.

AI Tools

AI tools refer to a suite of software and applications that utilize artificial intelligence techniques to perform tasks that would traditionally require human intelligence. These tools leverage machine learning algorithms, neural networks, natural language processing, and other AI technologies to automate processes, enhance decision-making, recognize patterns, interpret data, and facilitate complex tasks.

AI tools can be found in various domains including data analytics, image and voice recognition, chatbots, predictive modeling, and recommendation systems, among others. They are designed to improve efficiency, accuracy, and scalability, often adapting and evolving based on the data they process.

AI Tools

- a. Platforms and software available for teaching statistics
- b. Examples of activities and projects that use artificial intelligence
- c. Recommendations and best practices for effective use of artificial intelligence in teaching statistics

Examples

- SEIR Model for COVID-19
- Regression Analysis
- Conjoint Analysis
- Multivariable Analysis on Poverty
- Neural Analysis
- Bayesian Forecast for COVID-19
- Conjoint Analysis for Smartphone Development

Challenges and obstacles in implementing artificial intelligence in statistical education

```r
# Installing and loading the igraph package
install.packages("igraph")

library(igraph)

# Sample data: each pair represents two people (by their IDs) who interacted
edges <- data.frame(
  from = c(1, 2, 3, 4, 5, 6, 7, 8, 9),
  to   = c(2, 3, 4, 5, 6, 7, 8, 9, 10)
)

# Creating the graph
g <- graph_from_data_frame(edges, directed = FALSE)

# Setting labels for our nodes (people)
V(g)$name <- c("Chris", "Alex", "Jamie", "PersonD", "PersonE", "PersonF",
               "PersonG", "PersonH",
               "PersonI", "PersonJ")

# Plotting the network
plot(g, vertex.size=30, vertex.label.cex=0.8)

```

Figure: Network Analysis

Prompts for Education repository!

GITHUB

Administrators, teachers, and other staff members can utilize these prompts to

1. Create Engaging Lessons: Quickly design interesting and interactive lessons that captivate students.
2. Answer Student Questions: Provide accurate and fast answers to common student inquiries.
3. Automate Routine Tasks: Simplify day-to-day tasks with ready-to-use prompts.

How to Use

1. Find a Prompt: Browse through our collection (currently a work in progress).
2. Copy-Paste: Follow the direct link to Bing Chat or highlight, copy, and paste the prompt into your GPT-powered tool.
3. Apply the Answer: Use the response in your teaching, administrative tasks, or educational activities

AI's Impact on Employee Performance

Employees using ChatGPT outperformed those who didn't:

- Speed increased by 25.1 pp
- Output quality improved by over 40 pp
- 12,2 pp more tasks completed

AI has a "skill-leveling effect":

- Lower-performing consultants received a 43 pp performance boost
- Narrowed the performance gap between top and bottom consultants

More info

Conclusion

Challenges with AI's limitations:

- - Task outside AI's capability saw a performance decline (84pp to 60-70pp accuracy)

- Two effective strategies to harness AI:

- 1. Centaur Behavior: A distinct division between human and AI tasks.
- 2. Cyborg Behavior: Seamless integration, where AI and humans work hand-in-hand.

Over-reliance on AI can have pitfalls:

- - Reduced vigilance in decision-making
- - Homogeneity in AI outputs

More info

The Future

- **AI's transformative potential:**
  - - Significant enhancement in work productivity and quality
  - - Need to understand the 'Jagged Frontier' of AI for optimal benefits
- **Future implications:**
  - - As AI evolves (beyond GPT-4), organizations must adapt for ethical and effective integration
- **Accessibility:**
  - - Advanced AI tools, once exclusive, are now accessible to a wider audience
  - - Both elite consultants and the general public can leverage AI for improved work outcomes

Why not?

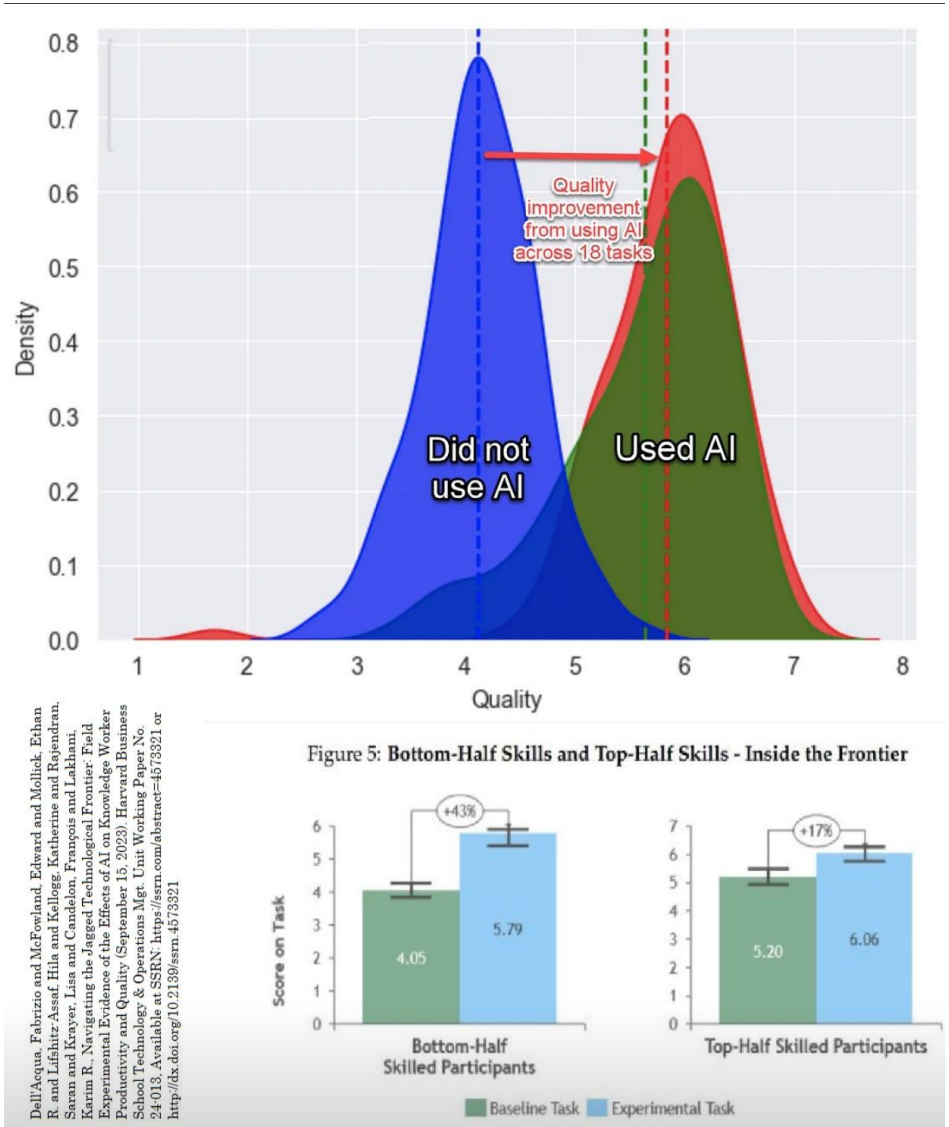


Figure: Use or Not use AI

Why not?

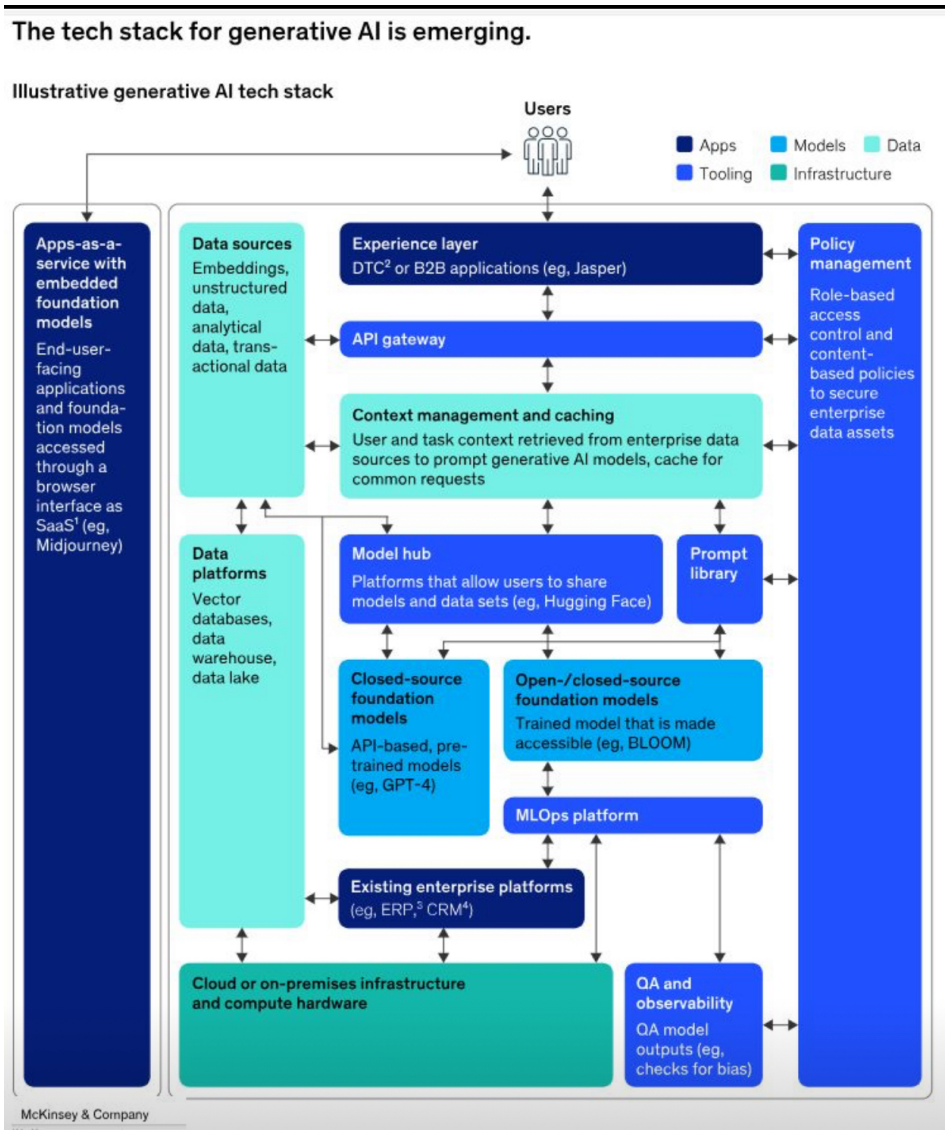


Figure: Tech stack for generative AI

From Books to Star Trek

Generate an image using BING that illustrates how AI bridges the gap between traditional books and the futuristic world of Star Trek, making it as futuristic as possible

From Books to Star Trek

Generate an image using BING that illustrates how AI bridges the gap between traditional books and the futuristic world of Star Trek, making it as futuristic as possible

Where to Star

You need to beggin here....

- Code that `https://code.org/ai/pl/101`
- Future Tools `https://www.futuretools.io/`
- Run my ... `https://app.runwayml.com/`
- PromptHero `https://prompthero.com/featured`
- FutureTools `https://www.futuretools.io`
- HugginFace `https://huggingface.co/`
- Sequoiacap `https://www.sequoiacap.com/`

Let's Talk

- Prof. Agustín Gómez Meléndez,. MSI.
- Oficina/Teams: +506 2511 5474
- Cel: 6003-5479 / 8317-5479
- Whastapp Institucional: 2511-4857
- Linkedin
- X
- Email: agustin.gomez@ucr.ac.cr / agomezme@uned.ac.cr

Let's Talk

Figure: Who will obtain the Degree?