Part 5: Generative AI

The goal of these slides is to review the previous week's content and to dive into a deeper analysis of the basic mechanics of Generative AI.

This document is designed to walk you through the content and to collect notes that might be helpful as you go.

About The Tool(s)

DALL-E is an artificial intelligence (AI) text-to-image model developed by Open AI. It can be used to create a new and unique image, based on a prompt given. Similar to Chat GPT, it uses reinforced learning to refine the quality of its searches.

Slideshow contents:

- Recap/Looking ahead
- Weekly reflections
- Review of Predictive Analytics and AI: What it does
- Generative AI Overview
 - o What is it?
 - o How does it work?
 - o Limits and Strengths of Generative AI
- Activity: Generative AI and Prompt Engineering
 - o Hand out activity guide
 - o Share results!
- Next Time:
 - o AI in our communities!

Last time, we:

- Shared the third round of reflective assignments
- Predictive Analytics & AI
 - o Definition
 - o The data it uses
 - o Where we see it today
- Demonstration: KNIME

Today will be about...

- Fourth round of reflective assignments
- "Generative AI"
 - o Definition
 - o The data it uses
 - o Activity: Prompt Engineering with Dall-E
- Debrief
- Look ahead at Next Time

Our Weekly Reflections:

• How it works (schedule)

Review: What Predictive Analytics And AI Do

• Go over the table

Generative AI

- Most used tools for Predictive Analytics
- Quick Intro to KNIME
- Example: Collecting and analyzing 10 years of NBA Data
- Can we predict player performance?

Tools for Predictive Analytics:

Two categories:

- Open Source
 - Python : several packages including TensorFlow (Google), PyTorch (Facebook)
 - R
 - KNIME

- Proprietary
 - Proprietary
 - IBM Watson Studio
 - Microsoft Azure Machine Learning
 - Google Cloud AI Platform
 - Amazon SageMaker
 - DataRobot
 - H2O.ai (also an Open-Source version)
 - Databricks

What is KNIME?

- KNIME is a drag and drop tool used for statistical analysis and machine learning.
- Imagine you have a bunch of ingredients (data) that you want to cook up into a delicious dish (insights).
- KNIME helps you do that by providing all sorts of tools and gadgets (nodes) that you can use to prepare, mix, and analyze your ingredients.
- You start with raw data, just like you start with raw ingredients for cooking. Then, you use KNIME to chop, dice, mix, and transform your data just like a chef would prep ingredients.
- Once everything is ready, you can use KNIME to bake, fry, or boil your data (analyze, visualize, and model) until you get the perfect result (insights).

KNIME for AI

(Demonstrates drag-and-drop function and assembly)

Descriptive Statistics

(Column = topics, including position, age, etc.)

KMEANS in KNIME

Imagine you have a big basket of colored marbles, but they're all mixed up. Your job is to group them based on their color. The KMeans algorithm is like a smart sorting machine that helps you do this automatically.

• Put the Marbles in the Machine

• Decide on Groups

• Sorting Begins

• Repeat Until Sorted

NBA Clustering in KNIME

- Big questions:
 - o Can we predict how good a player is?
 - o We don't have any label on our data
 - o We could try to sort all the players into two or three categories
 - o Which categories could we think off?
 - o How do we do that using KNIME?

• {Show clusters]

- Creating Rules based on the clusters
 - o Could a Tree Model help us predict if a player is outstanding?
 - o Why is the prediction not so straightforward?
 - o How much data is enough?
 - o What should we consider when using historical data?

What is a classification tree model?

You're planning a vacation but you're not sure where to go. You have a list of destinations and criteria that are important to you, like weather, activities, and cost. The KNIME tree model is like a super-smart travel advisor that helps you make the best decision based on your criteria.

• Destination Options

• Decision Criteria

• Branching Decisions

• Final Recommendations

[Halfway point] Last time, we...

- Shared our third round of weekly reflections
- Did a deep dive on Predictive Analytics & AI
- KNIME Activity
- Any questions? Refreshers?

Today, we will...

- Continue Predictive AI discussion
 - Terms
 - Group activity: Teachable Machine (Google)
 - Ethics
- Talk about the final project, etc.
- Debrief
- Look ahead at Week 5
 - Generative AI

Terms used last time

Note: feel free to integrate these notes and slides into the KNIME presentation itself. These were created after the fact.

- We heard some new terms during Monday (definitions on next slide):
 - o Normalization

o K-Means

o Decision Trees

o Random Forest

• Mostly related to programming side of AI

Activity: Teachable Machine

The activity is based around examining how data that is gathered can be used for further usage by AI programming.

Hand out worksheets.

The Ethics of Predictive Analytics and AI

- "Predictive AI" has many positive uses
- But even the most accurate data is a liability when it's misapplied.
- When we make mistakes with our predictions, what are the consequences and for whom?

- Let's look a few types of ethics:
 - o Unequal access
 - o Hacking and privacy
 - o Unanticipated impacts
- •

• "That brings risks"

Unequal Access

- **Sports**: Since Moneyball, teams have used predictive analytics to develop better lineups and strategies.
 - o But what happens when wealthy teams can do the same but with deeper pockets?
 - o Are these strategies always in the best interests of the players?

• **Municipalities**: Do some cities and towns have more resources to build predictive models for action?

• **Communities**: Are community members included in conversations about what we should use predictive AI to do and how we do it?

Hacking and Privacy

• **Data usage**: How is the data collected? Did I give consent?

• Access: Are any third parties looking at the data?

- **Retention**: How long will this data be held for?
 - o Policies vary (or may not even exist!)

Unanticipated Impacts

• Fire Department vs. Police Department dilemmas

• Predictive data cycle

Final Project

Edit as you see fit

- Break into groups (no more than three people) based on common community type (remember week one)
- Pick a situation/use case related to that community where AI could play a part.
- Define potential ethical limits and issues
- Present findings via a media form of your choosing (video, slideshow, etc.)
- In addition to your media presentation, write two paragraphs that summarize your project and explain how this issue also impacts your city, whether as a whole or in your respective neighborhood(s).

- The group's write-up ("memo") included:
 - o Background
 - Their community
 - AI and their community
 - o Options
 - Positive impacts
 - Potential ethical issues
 - o Recommendations
 - Important values for AI policies
 - Suggestions

Debrief and Upcoming reflections

- Debrief! A chance for you to tell us how it's going. For instance:
 - What went well this week?
 - What didn't?
 - What are you excited for?
 - What are you unsure about?

• PREDICTIVE ANALYTICS & AI: BONUS READING

Note: no weekly reflection was prepared for this week. However, a fitting substitute may be to have the participants present the fire department/police department scenario to their friends and families while responses are noted.

•