Defining Data Science

Brian Wright and Peter Alonzi

11/7/2022

Table of contents

Preface

1 Defining Data Science - 4 + 1 Overview

2 The 4 + 1 Model of Data Science

R.C. Alvarado, UVA School of Data Science

Data Science is a complex and evolving field, but most agree that it can be defined as the intersection of computer science and technology, math and statistics, and domain knowledge, with the purpose of extracting knowledge and value from data. Many also associate it with a series of practical activities ranging from the cleaning and "wrangling" of data, to its analysis and use to infer models, to the visual and rhetorical representation of results to stakeholders and decision-makers.

This essay proposes a model of data science that is intended to go beyond the laundry-list defintions that dominate the discourse in the field today. Although these are not inaccurate, they do not get at the specific nature of data science or help distinguish it from adjacent fields such as computer science and statistics — fields whose members sometimes claim to already be doing data science. Without a clear understanding of data science the field is subject to counterproductive turf battles in the academy as well as confusion in the workplace.

> `It is increasingly the case that hiring managers in industry understand the role of data s`

We define data science in terms of a multi-part model thats represents core areas of expertise in the field and how they are related to each other. These are the areas of **value**, **design**, **systems**, and **analytics**. A fifth area, **practice**, integrates the other four in specific contexts. Together, these areas belong to every data science project, even if they are often unconnected and siloed in the academy.

Unlike traditional academic disciplines, each area of the proposed model is inherently interdisciplinary, bringing together diverse and sometimes contrary perspectives under a common heading. The inherently interdisciplinary and pluralist nature of these areas is a distinctive feature of data science and a key differentiator between it and traditional disciplines.

The following describes how this model is derived and provides clues about how to interpret and apply the model to your own situation.

2.1 A Common Theme: The Image of the Pipeline

A review of the literature of data science definitions, from sources attempting to define the field explicitiy, as well as from self-definitions from adjacent fields such as data analysis and data mining, reveals that most definitions invoke the image of a data pipeline — a sequence of actions through which data flows as it moves from the consumption of raw data to production of results. Consumed data may come from a variety of sources — databases or intentional experiments or sensors. Results may be equally various, from the communication of analytical results to stake-holders to the development of a data product for use on the web. For a detailed review of these sources, see the Appendix.

```
The idea that this pipeline is fundamental to data science is not new &mdash; in the 1990s
```

2.1.1 A Common Sequence

An analysis of a representative sample of definitional essays shows that the various pipeline stories consist of elements drawn from a standard sequence of about twelve elements, give or take a few, depending on how one might expand or contract terms. These may signified by a core set of verbs or event types (which narratologists call functions), with the understanding that many synonyms are employed in the examples:

1. Understand
2. Plan
3. Collect
4. Store
5. Clean
6. Explore
7. Prepare
8. Model
9. Interpret
10. Communicate
11. Deploy
12. Reflect

No one definition includes them all, but some are more comprehensive than others, and different disciplines emphasize different parts.

For example, Hayashi's statistically-oriented definition of data science includes just three phases — design [1], collect [3], and analyze [8, 9] — with an emphasis on the experimental design phase in which data are actually produced through thoughtfully designed experiments (Hayashi 1998).

Mason and Wiggins propose five — obtain [3, 4], scrub [5], explore [6], model [8], and interpret [9] — which highlights two conditions that define industrial data science, the simultaneous availability of data (one merely obtains it, say through web scraping) and their poor condition relative to analysis, i.e. the need to scrub and wrangle them into usable form (Mason and Wiggins 2010).

The CRISP-DM model is the most comprehensive, with seven phases defined (if we include the unnamed but visually depicted function of storage), emphasizing the importance of understanding both the business proposition and data before anything is done with it (Wirth and Hipp 1999). It also modifies the metaphor of the pipeline, representing it as a circular and iterative process. However, unlike Donoho's similarly comprehensive sequence (implied by the ordering of his six divisions of "greater data science"), it does not include a "meta" phase devoted to reflecting on the process as a whole (Donoho 2017).

This twelve-part composite pipeline can be simplified by combining functions that naturally go together, by virtue of the expertise required to carry them out. This reduction yields about seven phases:

A understand and plan
B collect and store
C clean, prepare, and explore
D model and interpret
E communicate
F deploy
G reflect

Each of these may be considered a "chapter" in the story. Note that the number of verbs in each chapter title does not necessarily predict the length of its content. For example, the chapter on "model and interpret" covers a wide range of activities from a variety of perspectives, including classical statistics, machine learning, and computational simulation. It's a big and complicated chapter, but it is just one chapter among seven, even though many may consider it to be the most important chapter.

2.1.2 An Arc with Four Zones

`{figure} media/image1.png The Standard Sequence as a Narrative Arc`

To be sure, the middle chapter plays a central role in our story. If we think of the story as following a classical "there and back again" structure — a chiasmus pattern like X_1, Y_1, Z, Y_2, X_2 — then chapter D is the pivot, while chapters A, B, and C mirror E, F, and G. Thinking of the story in this way allows us to identify a parallel structure in the pipeline, connecting phases that are usually seen as separate. Specifically, we may visualize the pipeline as an arc, in which chapters in the first half of the pipeline mirror the those of the second half. We may then group chapters by the pairs formed in this way, yielding four zones — A and G belong to zone I, B and F to II, C and E to III, and D to IV — as in the following diagram:

`{figure} media/image2.png The Arc Transposed` With this visualization, we can discern some interesting properties about the data science pipeline that are not obvious in the original sequential image. For one, the arc structure suggests that the two ends of the pipe are not separate; both make direct contact with the external world. The external world — natural or social — from which data are pulled is the same world into which data products are inserted. This insight echoes the CRISP-DM model, which connects A and G (actually F), except that the two ends of the arc model are not directly connected. Instead, they come into contact with — and are separated by — the world in all of its complexity and unpredictability. The relationship between the effects caused by our data products G and the data we pull from the world A is not given but a matter of discovery — and often surprise.

At this point, we can explore the unifying themes associated with the four zones in our arc model by transposing the preceding visualization, which draws attention to what is common to each pairing. This generates four candidate areas of data science expertise — activities that, although they appear on opposite ends of the pipeline, nevertheless share basic knowledge, know-how, and areas of concern.

Zone IV is the easiest to interpret in this way because as the pivot of the arc it is not paired. It represents the work of modeling a problem mathematically, as well as evaluating and interpreting the results of mathematical modeling. This work requires data to be available in a particular form — clean and organized, usually as "tidy" analytical tables.

Zone I is also relatively easy to interpret: the functions in this group each involve understanding the relationship between the pipeline and the external world, the messy interface between the enterprise of data science and the variety of real world situations in which it operates.

We may note in passing that I and IV can be contrasted in several ways — messy vs clean, exoteric vs esoteric, qualitative vs quantitative, existential vs essential, concrete vs abstract, etc.

When it comes to zones II and III, the interpretation of results is less straightforward. This is because the reality of the kind of work performed in these areas is not as clear-cut as it is for I and IV. Both II and III exhibit an internal complexity not found in the others, and the two are less clearly separable from each other than they are from the other two. One reason for this complexity is that here pure and applied forms of knowledge intermingle in ways that defy easy description from an academic perspective.

For example, the work of "data wrangling," considered distinctive to data science, spans the two domains and involves a complex mixture of specific technological know-how and general scientific principles. It turns out that the relationship between these kinds of knowledge is highly contested, as evidenced by the reception of Donoho's "50 Years of Data Science," which has been criticized for separating science from engineering and demoting the importance of the latter (Donoho 2017). Regardless of the validity of this criticism, there is without doubt a long-standing conflict between data mining and data analysis over what counts as valid forms of knowledge, and this conflict emerges in the representation of zones II and III we find in our corpus.

We can take the conflict of interpretations over the status of technical knowledge in data science as a clue and use it to identify two broad dimensions that cross-cut the functions in zones II and III: technical know-how and abstract representation. Technical know-how II'

involves expertise in developing and deploying software and hardware designed to handle data at scale, including high-performance computing, big data architectures (such as Hadoop and its descendants), and data-oriented programming languages and libraries.

The topics associated with II' are highly specific and change rapidly relative to other forms of knowledge, and so are often omitted from, or under-represented in, academic curricula, even though to many they are the *sine qua non* of data science. Abstract representation $III´$, on the other hand, involves expertise in areas ranging from how data are to be modeled for capture and analysis to how the results of analyses are to be presented to non-expert decision-makers. These areas of knowledge strive for formal generality over the long run; they are often expressed as grammars or design languages, frequently with visual modes (such as entity-relationship models and unified modeling language UML). They also include other forms of visualization, such as the plots developed for exploratory data analysis, such as box plots, and those used to represent statistical facts and analytical results in dashboards and infographics.

2.1.3 The Four Areas, Plus One

We are now ready to define and name the areas of data science expertise that emerge from an analysis of the pipeline considered as an arc. In each case, we want to identify the common context shared by the paired activities in each zone as well as the tension that exists between them by virtue of their occupying opposite sides of the pipeline. In many cases, although we can identify a shared theme in each zone's work, the reality is that practitioners do not always interact or share disciplinary homes. One of the benefits of this model will be to identify these points of synergy and to identify new disciplinary boundaries.

Area I: Value

The area of value is defined by the relationship of data science to the world from which it draws data and into which it inserts data products. More broadly, it concerns the primary motivations of data science — why do we practice data science in the first place? It combines the traditional discipline of ethics with the professional activities of business planning, policy making, developing motivations for scientific research, and other activities that have a direct impact on people and the planet. This is the area where we determine what we do versus what we do not do, in order to maximize societal and environmental benefit and minimize harm. It is also the area that looks inward to the other data science areas and provides guidance on such issues as algorithmic bias or open science. Common activities include the forming of value propositions that initiate data science projects, research into how data is created and used "in the wild," understanding the ethics of data acquisition, manipulation, communication, and sharing, and the application of data products in the world.

Area II´: Design

The area of design is defined by the relationship between human and machine forms of representation. This relationship is bidirectional: human-generated data flowing into the pipeline must be represented for machine consumption (H2M, or $H \rightarrow M$), while analytically transformed data going out must be represented for human consumption (M2H, or $M \rightarrow H$). This area therefore includes expertise in human-machine interaction as it appears at the points of both consuming data and producing data products. Activities here include the representation and communication of captured data for the work of analytics, e.g. in database modeling, the curation of data, and of complex data and analytical results to humans to drive decision-making and influence behavior. It also includes the making of things, with purpose (i.e. to solve problems) and intent (meaning, concision, focus). A key part of the area is the broad practice of what is often called visualization, the translation of complex quantitative information into visual (and other sensory) forms that non-experts can understand. In slightly more technical terms, the area of design focuses on what Zuboff called "informating," the process by which the world is represented for computation and analytics, and also by which analytical models and results are represented to the world (Zuboff 1995). These two processes often produce competing representations — a private one *of* the world for the data scientist, and a public one *for* the world of the results of analytics. One task of this area is to reconcile these two representations.

Area III´: Systems

The area of systems is defined by the technological infrastructure that is common to the pipeline but concentrated in the activities of wrangling data, deploying data products, and building out systems to support these activities at scale. This area includes expertise in infrastructure systems and architectures to support working with big data — big in terms of volume, velocity, and variety — and building high performance systems in both development and production environments. It includes the broad areas of hardware and software as such — computer technology as opposed to computer science. Key activities include developing cloud resources, building performant pipelines to ingest and aggregate data, developing networks of resilient distributed data, and writing and using software to accomplish tasks. This area is often referred to as "data engineering" or "machine learning engineering," which, according to Owen, "is most of what Data Science is and Statistics is not" (Owen 2015).

Area IV: Analytics

The area of analytics is defined by the practice of mathematical modeling based on data. This area includes what many consider to be the essence of data science, the combination of statistical methods with machine learning, along with information theory, optimization, network analysis, complexity theory, simulations, and other rigorous quantitative methods from a variety of fields. Although unified by a broad commitment to advanced mathematical models and computational algorithms, in reality this is a heterogeneous collection of competing schools and methods. Tensions include inference vs prediction, parametric vs non-parametric (kernel-based) methods, frequentist vs Bayesian statistics, analytic vs algorithmic solutions (including simulations), etc. Key activities include clustering, pattern recognition, regression, rule mining, feature engineering, model selection, performance evaluation, and a host of other activities. Although currently dominated by statistical methods, this area also includes the rule-based methods that dominated the field of artificial intelligence before the more recent successes of statistical learning and deep learning.

Area V: Practice

The preceding four areas each represent areas of foundational knowledge, forms of expertise that can be taught as more or less separate subjects. In practice, however, these areas represent the interlocking parts of a division of labor that are integrated in the pipeline. This area consists of actual activities that brings people together to combine expertise from each of the four areas. It is characterized by data science teams working together and with external parties to develop solutions and projects that are responsible, authentic, efficient, and effective. Practice is also where the core areas of data science come into contact with a broad spectrum of domain knowledge and real world problems. The following diagram () shows the central, integrative role played by practice:

`{figure} media/image3.png The Integrative Role of Practice` ### Two Principal Components

Is there a way to understand how the four primary areas are related to each other, beyond their being composed of functions from the same pipeline? Put another way, does the pipeline-as-arc model exhibit any structural features that will help us conceptualize the broader space of data science? Two such features stand out: (1) the opposition between concrete and abstract forms of representation, and (2) between human and machine processing.

Regarding the concrete and the abstract, it's clear that the arc model has a metric quality to it: as one moves toward the pivot point of analysis, one moves away from the concrete messiness of reality as experienced to the "tidy" and abstract world of mathematics; similarly, as one moves from the pivot back to the world, there is a requirement to convert esoteric results into more humanly intelligible forms, often through a process of concretization; visualizations succeed by employing concrete metaphors that flesh out mathematical ideas that are notoriously detached from the imagination — no one can imagine, for example, n-dimensional spaces beyond a handful of dimensions. The arc describes a dialectic of abstraction and concretization that defines the ebb and flow and data science work.

`{figure} media/image4.png The Four Areas in Two Dimensions`

The dimension of human and machine processing exhibits a similar duality, that between the conversion of information from humanly accessible forms, such as given by data acquired by instruments, into machine readable and processible forms, and the reverse. The process of moving from human to machine representations is a large part of what data capture, modeling, and wrangling is all about, while the process of converting the results of machine learning, broadly conceived, into humanly actionable form is what visualization and productization are all about. The reality of this dualism is captured by the concept of human-computer interaction (HCI), an established field that is applicable to both sides of the arc.

How do the four fundamental areas map onto these two dimensions? We can define each area as a combination of one pole from each duality; the four areas result from all possible permutations of the two dimensions. This produces the following high level characterizations of each area: (1) Value is concerned with concrete humanity, (2) Design with abstract humanity, (3) Analytics with abstract machinery, and (4) Systems is concerned with concrete machinery. All of these make intuitive sense, with the exception of Design. This is consistent, however, with the fact that the area of Design emerges from this analysis as an undervalued and not well understood area of expertise, even though Yau emphasized it early on (Yau 2009b). Indeed, one of the consequences of this analysis is to train our attention on this area of knowledge and to develop it further.

```
It is worth noting that the four combinations are surprisingly analogous to the four appro

(1) thinking humanly \
(2) thinking rationally \
(3) acting humanly \
(4) acting rationally

Moreover, it is easy to see how the following analogies make sense:

$$
abstract : concrete :: thinking : action
$$

and

$$
human : rational :: human : machine
$$


In fact, it appears that the same space is shared by the 4+1 model of data science and Rus
```

One exciting interpretation of the two dimensions defined here is that they correspond to two principal components that undergird the general field of data science. As components, these axes define two orthogonal dimensions within which all the specific topics of data science may, in principle, be plotted. The reality behind these axes may be that they represent cognitive styles associated with the division of labor implied by the data science pipeline.

PC1: Human versus Machine

The human-machine axis accounts for the most variance in the field. This seems evident from the fact that Conway's Venn diagram model of data science represents only the machine side of our model (with practice replaced by "substantive expertise"). The human side —

Value and Design — is left out, or short-changed by being lumped in with domain knowledge. The very fact that the human side has to be explained and added to the model suggests strongly that it defines a pole at some distance from the areas of knowledge described in Conway's model. The human pole refers to humanity understood as situated in their historical, social, and cultural milieu. It is synonymous with *human experience*. The machine pole refers to the technoscientific apparatus of formal, quantitative reasoning that operates on representations of the human and the world. In the context of data science, it is more or less synonymous with *machine intelligence*, broadly conceived to include machine learning but also other modes of analysis on the spectrum of prediction and inference. Given these poles, the human-machine axis represents the opposition between humanistic disciplines that seek to understand human experience as such, and the formal sciences that employ machine intelligence, broadly conceived, to interpret that experience as represented and aggregated in the form of data.

PC2: Concrete versus Abstract

The abstract-concrete axis accounts for the difference between two forms of knowledge, roughly between direct experience and the indirect representation of that experience enabled through data. Both the realm of Value and Systems involve immersion in the messy details of lived experience — and direct acquaintance with the devils in those details. This is the messy world of hacks and ironies. The realms of Design and Analysis, on the other hand, are founded on abstract representations that strive for clear and distinct purity, and which allow for deductive reasoning to succeed at the cost of simplifying assumptions and reduced representations. This is the orderly world of models. The concrete pole refers to situated knowledge, knowledge as understood by hackers and makers, but also ethnographers who seek to maximize thick description in their work. It represents *concrete materiality*. The abstract pole refers to formal knowledge, knowledge in the form of mathematical symbolism, deductive proofs, and algorithmic patterns. It is *abstract form*. Given these poles, the concrete-abstract axis is roughly the opposition between applied and pure forms of knowledge, between those that embrace materiality and those that seek purity of form.

2.2 Final Representation

The result of the preceding may be represented by the following graphic.

`{figure} media/image5.png The 4+1 Model of Data Science`

This visualization represents data science as composed of specific and complementary forms of knowledge. The vertical axis defines the dominant polarity between analysis — the *how* of data science, often identified entirely with it, contrasted with the *why* of data science, from which data science derives its meaning and value as a profession. The horizontal access defines the polarity of methods that are often obscured in academic definitions of data science — the supporting practices that make the Analytics component work in the first place.

2.3 Concluding Remarks

The point of the $4 + 1$ model, abstract as it is, is to provide a practical template for strategically planning the various elements of a school of data science. To serve as an effective template, a model must be general. But generality if often purchased at the cost of intuitive understanding. The following caveats may help make sense of the model when considering its usefulness when applied to various concrete activities.

The model describes areas of academic expertise, not objective reality. It is a map of a division of labor writ large. Although each of the areas has clear connections to the others, the question to ask when deciding where an activity belongs is: *who would be an expert at doing it*? The realms help refine this question: the analytics area, for example, contains people who are good at working with abstract machinery. The four areas have the virtue of isolating intuitively correct communities of expertise. For example, people who are great at data product design may not know the esoteric depths of machine learning, and that adepts at machine learning are not usually experts in understanding human society and normative culture.

Each area in the model contains a collection of subfields that need to be teased out. Some areas will have more subfields than others. Although some areas may be smaller than others in terms of number of experts (faculty) and courses, each area has a major impact on the overall practice of data science and the quality of an academic program's activities. In addition, these subfields are in an important sense "more real" than the categories. We can imagine them forming a dense network in which the areas define communities with centroids, and which are more interconnected than the clean-cut image of the model implies.

The principal components abstract/concrete and human/machine are meant to help imagine the kinds of activities that belong in each area, through their connotations when combined to form the four bigrams — concrete human, abstract human, concrete machine, and abstract machine. For example, the area of value as the realm of the "concrete human" (or perhaps "concrete humanity") is meant to connote what the Spanish philosopher Unamuno called the world of "flesh and bone" within which we live and die, that is, where things matter. On the other hand, analytics as the realm of the "abstract machine" is meant to connote the platonic world of mathematical reasoning which, since Euclid, has been characterized by rigorous, abstract, deductive reasoning that has literally been described as an abstract machine (see Alan Turing).

At the center of this model and each area is people. Even in the area classified as "abstract machine," people and human thinking is at the center.

Part I

Design

This section will provide a overview of the the Design space through three different lenses. The School of Data Science's perspective on the curriculum and skills that compose this area, the view from industry to included recorded lectures from experts that have participated in the Foundations of Data Science Course at UVA and through the point of view of researchers that are working and contributing to this area of the field.

It is important to note that all these areas are evolving and have overlapping content area, but Design in particular can be difficult to neatly fold into well-defined set of easily quantifiable skills.

3 Design SDS

Overview of the Curriculum and Skills that the School of Data Science and UVA believes to be included in the domain area of Design.

4 Design Lab

Overview of the Design lab for the current semester for Foundations of Data Science Course at UVA.

5 Design External to the School

Contains a overview of the perspective from industry or others outside of academia on the Design space of Data Science.

6 Design Case Study

Overview of the Design case study for the current semester for the Foundations of Data Science course.

Part II

Value

7 Value SDS

8 Labs Rubric - Guess Who

DS 1001 - Spring 2023 - Professors Wright and Alonzi Due: End of lab period (or later that day) Submission format: Word doc or PDF summarizing your findings

Individual Assignment

General Descripition: This lab is design for you to reflect on the process used for winning the game Guess Who. Students should work in teams to play the game and take notes on the processes that work best for quickly guessing the opponents card. This will likely include the creation of a tree based decision diagram, a example will be provided in class.

Preparatory Assignments - None

Why am I doing this? In order to allow you to reflect on how data is used to generated decision/predictive algorithms and what issues could arise as a result. Work to incorporate the materials you've been exposed to up to this point as you work through the lab.

What am I going to do? You are working through a interactive process with your group to find the optimal list of yes/no questions for any given card. Play the game as many times as necessary to construct a list of the top five questions using a tree based diagram.

Answer these questions:

1. In this situation what is the dataset and what is the algorithm?
2. How did your approach to identifying cards change throughout the lab period?
3. Are there cards that are easier to identify, why?
4. This is just a game, but given the materials you've been exposed to what concerns do you have about this process and algorithmic decision making?

Tips for success:

- Don't worry about winning
- Work as a team with your group
- Try to document the process used right at the start of the lab, don't wait till after you have played to start taking notes

How will I know I have succeeded

Specs Category	Specs Details
Formatting	• Submit Via Canvas
Text	• Text answers to questions • Goal: The questions are designed to be answer during or right after the lab period. • Use a few sentences to answer each question in Canvas

Acknowledgements: Special thanks for Jess Taggart from UVA CTE for coaching us. This structure is pulled directory from Steifer & Palmer (2020).

9 Labs Rubric - AI Fairness 360

DS 1001 - Spring 2023 - Professors Wright and Alonzi Due: End of lab period (or later that day) Submission format: Word doc or PDF summarizing your findings

Individual Assignment

General Descripition: This lab is designed for you to get exposure to AI fairness approaches in a no code environment on the website AI Fairness 360. You will be able to work through the various fairness methods at different stages of the pipeline and reflect on which methods seem to work the best on the given datasets.

Preparatory Assignments - None

Why am I doing this? In order to give you exposure to and practice with the various methods being developed and deployed in the ML fairness space. After completing the lab you'll have a better sense of how these tools are used, when they are used and how the work.

What am I going to do? The AI Fairness 360 website has a demo module that includes three datasets. Work through the demo on all three datasets, trying all the methods provided, and answer the questions below.

Answer these questions:

1. For each protected class variable which evaluation methods showed bias to be present?
2. Note how each method preform at removing bias.
3. Was the accuracy of the model effected when using the various approaches, if so how?
4. Given the above what are some patterns you noticed, which methods seem to work the best, where in the data process are these methods located (pre/in/post).

Tips for success:

- Take careful notes as you go through each method
- Have fun

How will I know I have succeeded:

Specs Category	Specs Details
Formatting	• Submit Via Canvas • Text answers to questions
Text	• Goal: The questions are designed to be answer during or right after the lab period. • Bullett points are fine for the first three questions, paragraph from for the fourth

Acknowledgements: Special thanks for Jess Taggart from UVA CTE for coaching us. This structure is pulled directory from Steifer & Palmer (2020).

10 Labs Rubric - Explainable AI Interview

DS 1001 - Spring 2023 - Professors Wright and Alonzi Due: End of lab period (or later that day) Submission format: Word doc or PDF summarizing your findings

Individual Assignment

General Descripition: This lab is designed for you to verbalize the concept of Explainable AI to someone that is unfamiliar with the concepts and track their responses.

Preparatory Assignments - None

Why am I doing this? This will allow you to work on communicating the concepts thus understanding them better and judging whether trust is increased with the knowledge that explainable approach to AI models are present.

What am I going to do? Interview three people that have somewhat limit knowledge of Machine Learning and AI. Ask the first two questions then provide a brief overview of Explainable AI and then ask the remaining three questions. After all three interviews are complete provide a brief summary of what patterns you notice (a paragraph or two is fine)

Answer these questions:

1. Do you believe that AI has a large influence over society currently?
2. On a scale from 1 to 10, with 10 being total trust and 1 being no trust: Do you believe that AI models can be trusted and used in a way that promotes positive outcomes for individuals and society?
3. Give a brief overview of the Data Ethic concepts with a emphasis on Explainable AI models.
4. Given what you just heard does that in anyway change your answer to question 2, such that, if explainable AI models were more commonly used would you be more likely to believe that they can promote positive outcomes for individuals and society? (scale of 1 to 10 for you belief the AI models can promote positive outcomes)
5. Are you interested in learning more about Explainable AI as a concept?

Tips for success:

- Keep your over of data Ethics and Explainable AI at a high level and use the slides provided in class as needed.
- Feel free to improvise and ask additional questions if you'd like or change the questions slightly.

How will I know I have succeeded:

Specs Category	Specs Details
Formatting	• Submit Via Canvas • Summary of your three interviews in one document • A reflection on any patterns you noticed in the answers
Text	• Goal: The interview is designed to be fairly quick, 10 minutes or so for each. All together the assignment should take approximately one hour.

Acknowledgements: Special thanks for Jess Taggart from UVA CTE for coaching us. This structure is pulled directory from Steifer & Palmer (2020).

11 Value External to the School

12 Value Case Study

of this Value domain

Part III

Systems

13 Systems UVA DS

14 Systems Lab

15 System External

16 Systems Case Study

Part IV

Analytics

17 Analytics UVA DS

```
import numpy as np
import matplotlib.pyplot as plt

r=np.arange(0,2,.01)
theta = 2 * np.pi * r
fig, ax = plt.subplots(
  subplot_kw = {'projection': 'polar'}
)
ax.plot(theta,r)
ax.set_rticks([.5, 1, 1.5, 2])
ax.grid(True)
plt.show()
```

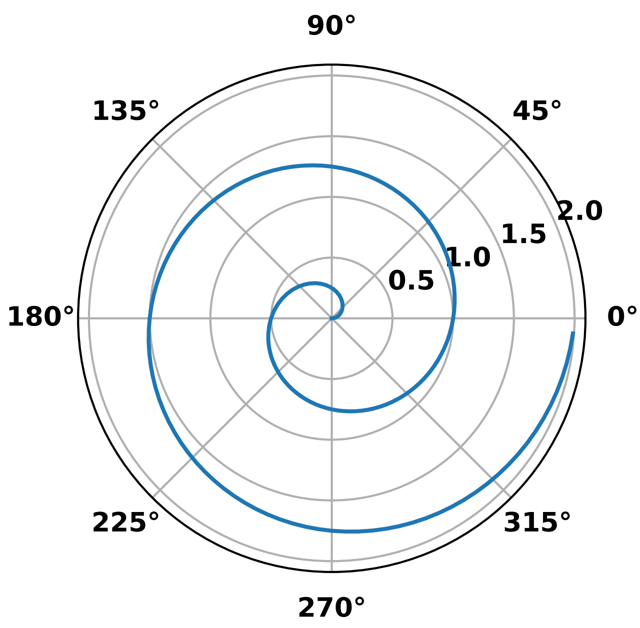

Figure 17.1: A line plot on a polar axis

18

Labs Rubric - Mastermind - Entropy

DS 1001 - Spring 2023 - Professors Wright and Alonzi

Due: End of lab period (or later that day)

Submission format: Word doc or PDF summarizing your findings

Individual Assignment

General Descripition: This lab is designed to introduce the basic concepts behind information gain and entropy. Mastermind is a game were through a series of questions one player tries to determine a "code" created by another player. In doing so the value of asking questions that provide a high level of information will become paramount.

Preparatory Assignments - None

Why am I doing this? In order to gain a better understanding of information gain and entropy through hands on experience.

What am I going to do? Work with a partner or group of three and alternate playing the game. At the onset of each game the code-maker should calculate the entropy of the chosen pegs for the "code". The code-breaker should note the results of the game . This should include whether the code was broken or not and the number of rounds used. Try to be the code-maker and breaker at least 3 times. The code-maker should intentionally use very different types of "code", this should impact how the game is played.

- Note the range of entropy for binary cases is 0 to 1 for more than 2 classes it is 0 to $\log_2 k$, where k is the number of classes.

Answer these questions:

1. What combinations of pegs (code) seemed to be harder to break?
2. Did your approach to asking questions change as you played?
3. Describe where in the game information gain is being presented?

Tips for success:

- Don't worry about winning, instead think about what is happening during game play.
- Work as a team with your group.

- Try to document the process used right at the start of the lab, don't wait till after you have played to start taking notes.

How will I know I have succeeded:

Specs Category	Specs Detail
Submission	Submit via Canvas a PDF or Word Document
Text	* Answer the above questions * When you were the code-maker submit the Entropy * When you were the code-breaker submit the results

Acknowledgements: Special thanks for Jess Taggart from UVA CTE for coaching us. This structure is pulled directory from Steifer & Palmer (2020).

19 Analytics External

20 Analytics Case Study

21 BSDS-Course-Info

- DS 2006
- DS 3005 Mathematics for Data Science

References