Journal of Machine Learning Research 21 (2020) 1-39 Submitted 10/19; Revised 3/20; Published 9/20

Rationally Inattentive Inverse Reinforcement Learning Explains YouTube Commenting Behavior

William Hoiles WILLIAM.HOILES@KATERRA.COM
KATERRA
Menlo Park, CA 94025*, USA*

Vikram Krishnamurthy VIKRAMK@CORNELL.EDU
Electrical and Computer Engineering
Cornell University
Ithaca, NY 14853*, USA*

Kunal Pattanayak KP487@CORNELL.EDU
Electrical and Computer Engineering
Cornell University
Ithaca, NY 14853*, USA*

Editor: Shie Mannor

Abstract

We consider a novel application of inverse reinforcement learning with behavioral economics constraints to model, learn and predict the commenting behavior of YouTube viewers. Each group of users is modeled as a rationally inattentive Bayesian agent which solves a contextual bandit problem. Our methodology integrates three key components. First, to identify distinct commenting patterns, we use deep embedded clustering to estimate framing information (essential extrinsic features) that clusters users into distinct groups. Second, we present an inverse reinforcement learning algorithm that uses Bayesian revealed preferences to test for rationality: does there exist a utility function that rationalizes the given data, and if yes, can it be used to predict commenting behavior? Finally, we impose behavioral economics constraints stemming from rational inattention to characterize the attention span of groups of users. The test imposes a Rényi mutual information cost constraint which impacts how the agent can select attention strategies to maximize their expected utility. After a careful analysis of a massive YouTube dataset, our surprising result is that in most YouTube user groups, the commenting behavior is consistent with optimizing a Bayesian utility with rationally inattentive constraints. The paper also highlights how the rational inattention model can accurately predict commenting behavior. The massive YouTube dataset and analysis used in this paper are available on GitHub and completely reproducible.

Keywords: Inverse Reinforcement Learning, Bayesian Revealed Preference, YouTube, Rational Inattention, Rényi Mutual Information, Framing, Behavioral Economics, Deep Embedded Clustering, Contextual Bandits

1. Introduction

This paper considers a novel application of inverse reinforcement learning with behavioral economics constraints to model, learn and predict the commenting behavior of YouTube viewers. We model each group of users as a Bayesian rationally inattentive agent. (Equivalently, each agent solves a contextual bandit problem with behavioral economics constraints.) Our methodology integrates three

key components to model the collective commenting behavior of YouTube viewers. First, to identify distinct commenting patterns over YouTube videos, we use *Deep Embedded Clustering* that partitions groups of videos from a massive YouTube dataset into non-overlapping segments. For each segment, we define a group comprising the users that view and comment on videos within that segment- each group's behavior is individually analyzed. The second component involves inverse reinforcement learning; we use *Bayesian Revealed Preferences* to construct a test for utility maximization behavior for individual groups of users and estimate utility functions that rationalize their behavior. Finally, we impose behavioral economics constraints stemming from *Rational Inattention* to characterize the attention span used by the viewers while commenting. This rationally inattentive model of commenting behavior can be interpreted as a constrained contextual bandit problem as will be discussed in Sec. 2. After a careful analysis of a massive YouTube dataset, our surprising result is that in most YouTube user groups, the overall commenting behavior[1] is consistent with optimizing Bayesian utility with rationally inattentive constraints. The paper also highlights how the rational inattention model can predict commenting behavior.

To better explain the main ideas, consider the following abstract setup: suppose a Bayesian agent chooses an action to maximize an expected utility function based on the noisy measurement of an underlying state. Assume that the Bayesian agent is rationally inattentive, that is, obtaining this noisy measurement is expensive – this information acquisition cost affects the action chosen by the agent. An observer (analyst) records the dataset of actions of the Bayesian agent and knows the underlying state. Classical inverse reinforcement learning aims to estimate the utility function of a decision process by observing its decisions and stimulus input (Ng et al. (2000)) while assuming the agent is a utility maximizer. Fu et al. (2017) discuss construction of disentangled state dependent rewards for agents. In such problems, the existence of a utility function the agent maximizes is assumed implicitly. The revealed preference approach in this paper addresses a deeper and more fundamental question: does a utility function exist that rationalizes the given data (with rational inattention constraints) and if yes, how to estimate this utility function.

Estimating utility functions given a finite length time series of decisions and budget constraints is widely studied in the area of revealed preferences in economics, starting with the paper of Afriat (1967) which gives a remarkable necessary and sufficient condition for utility maximization; see also Varian (1982, 2012); Woodford (2012) and more recently in machine learning (Lopes et al. (2009)). Varian (1983) characterized investor behavior by devising Afriat-type conditions for expected utility maximization in a Bayesian framework. Unlike Afriat (1967) and Varian (1983), the utility function in our Bayesian set-up involves discrete valued variables.

Costly information acquisition by Bayesian agents has been studied by economists and psychologists under the area of "rational inattention" pioneered by Nobel Laureate Christopher Sims (Sims (2003, 2010)). Rational Inattention is a form of bounded rationality - the key idea is that human attention spans for information acquisition are limited and can be modeled in information theoretic terms as a Shannon capacity limited communication channel. Sim's rational inattention model is studied extensively in behavioral economics (Matejka and McKay (2015)). Woodford (2012) considered an upper bound on the Shannon capacity for testing rational inattention with visual perception queues. Typically, the information acquisition costs faced by a decision maker are not known to the observer (analyst). A general test for rational inattention is proposed in two remarkable

1. By overall commenting behavior in YouTube, we mean both comment count and video ratings (likes and dislikes). Another term used in the literature (Khan (2017b)) is "user engagement". This paper characterizes the overall user engagement in a YouTube dataset.

papers by Caplin and Martin (2015) and Caplin and Dean (2015) with minimal restrictions on the information acquisition cost- we will refer to this as the general cost.

Framing and Categories

Our analysis involves partitioning the massive YouTube dataset into segments using two different methodologies. The first method constructs non-overlapping segments of videos based on the nature of their framing information gathered from the videos' thumbnail and description. In the second method, the videos are grouped based on the video category (one of the attributes in YouTube data)- all videos falling under the same category constitute a single segment. We shall use the term "frame" and "category" to refer to segments when analyzing data partitioned in the first and second way respectively.

Let us briefly explain these two methods: In behavioral economics, Tversky and Kahneman (1981) use "frames" to describe information an agent has when making a decision. Regarding analysis of YouTube and other social media datasets, extensive studies (Khan (2017a); Kwon and Gruzd (2017); Alhabash et al. (2015); Hoiles et al. (2017)) show that comments posted by users are influenced by the thumbnail, title, category, and perceived popularity of each video. Lange (2007) shows that sharing and circulation of videos is a manifestation of social relationships among youth, thus linking human cognitive behavior to user engagement dynamics. For example, when selecting which product to purchase on a website, the positioning of the products and surrounding content on the website impacts how humans select a product. Given such external information (image/text/numeric) in which the decision problem is embedded, how can one construct a tractable feature set? We develop deep embedded clustering methods to construct the frames to test for rational inattentive agents. The deep embedded clustering is based on Xie et al. (2016); Guo et al. (2017), however we design the input, encoder, and decoder to account for the visual perception of the frame of the decision problem which includes image, text, and numeric information.

Commenting behavior in YouTube has been well-studied in literature. Siersdorfer et al. (2010) analyze comments and their ratings in various categories of YouTube videos. They illustrate the increasing variance of comment ratings and sentiments with polarizing content in different categories of videos. Using open source sentiment analysis tools, the authors also predict the anticipated comment ratings (community feedback) on unrated comments in YouTube videos. Schultes et al. (2013) define comment classes to analyze YouTube comments on videos over various categories. They conclude that the distribution of comments over comment classes differs with video category. Severyn et al. (2014) predict YouTube comment types by classifying them using existing comments in contrast to our case, where we predict the commenting behavior using the videos' content.

To summarize, the two methodologies we consider, namely frame and category, constitute two distinct ways of partitioning groups of videos in YouTube data - a *user-driven approach* based on how the user infers the videos and a *content-driven approach* based on the video category. The perceived popularity of the underlying state is indicated by each video's viewcount, and the nature of commenting behavior is characterized by the number and sentiment (captured by likes and dislikes for each video) of the comments. Hoiles et al. (2017) state that first day viewcount, number of subscribers, contrast of the video thumbnail, Google hits, and number of keywords are the five dominant meta-features that affect the popularity of a video. However, in this paper, we consider only the first meta-feature as a popularity indicator for simplifying our YouTube analysis and emphasize

more on the information acquisition procedure of the YouTube viewers before a commenting action takes place.

The two significant extensions considered in this paper are the effects of framing (determined using deep embedded clustering) and video category and the use of Rényi mutual information cost for testing rational inattention. Our rational inattention test is equivalent to solving the temporal credit assignment problem in *preference-based inverse reinforcement learning* (Wirth et al. (2017)). Such inverse reinforcement learning is used with non-numeric feedback (Wirth et al. (2016)), e.g. in socially adaptive path planning (Kim and Pineau (2016); Hadfield-Menell et al. (2017)) for robots. As will be discussed in Sec. 2, the methodology in this paper is equivalent to inverse reinforcement learning of a contextual bandit with rational inattention (behavioral economics) constraints.

Main Conclusions on YouTube Dataset

This paper focuses on an application of inverse reinforcement learning with behavioral economics constraints to predict overall YouTube commenting behavior. So to better motivate the paper, we summarize our main findings (see Sec. 6 for details). Based on extensive data analysis on groups of YouTube users, where each group consists of approximately 3500 viewers, our main take-home message (from a behavioral economics point of view) is that YouTube user groups are rationally inattentive in their commenting behavior and that users prefer to comment on videos that are perceived to be popular. In YouTube, video viewcount is the independent quantity which governs the commenting behavior since videos need to be viewed first before users can comment or rate the video. Our analysis thus also sheds light on the way videos are "perceived" by groups in YouTube i.e. how does the collective sentiment behind popular and non-popular videos vary over different video segments. The set valued utility estimates of commenting behavior for user groups also help in predicting future behavior of the users in each video segment; see Sec. 6 for additional conclusions. Our pre-processing uses state-of-the-art NLP tools like GloVe (Pennington et al. (2014)) to extract information from videos' text descriptions and combines it with *Deep Embedded Clustering* to process videos' metadata for partitioning the YouTube dataset. The results presented in Sec. 6 are completely reproducible, with the code and datasets publicly available on a GitHub repository.

Why Analyze the Commenting Behavior in YouTube?

YouTube is clearly a social media site, however, YouTube is also a social networking site. Classical online social networks are dominated by user-user interactions. However YouTube is unique in that the interaction between users includes video content–that is, the interaction follows users-content-users. Examples where our analysis of YouTube commenting behavior are relevant include:

1. *Behavioral Psychology*- Understanding and modeling how humans perceive social multimedia information in the context of framing and rational inattention (Poggi and D'Errico (2010); Tutino (2011)).
2. *Economics of YouTube*- Knowledge of commenting behavior characteristics allows YouTube partner companies to adapt their user engagement strategies to generate higher viewcounts and increase revenue (Hoiles et al. (2017)).
3. *Content Caching in* $5G$- By caching highly popular content at base stations (BS), user demand for the social media video content can be served locally. This reduces the overall network traffic and improves the QoE for users requesting content (see Hoiles et al. (2015) and references therein).

4. *YouTube Recommendation*- The variance of the utility function with respect to the optimal action selection policy is a useful measure of uncertainty involved with a Bayesian decision maker (Levy and Markowitz (1979); Mannor and Tsitsiklis (2011)). For the YouTube dataset, we found that this variance is higher for categories like Sports, News and lower for less controversial categories like Automobiles, Education. These are in line with results in Siersdorfer et al. (2010) which reports a higher variance of comment ratings in polarizing topics compared to neutral topics. This variance appears in our finite sample analysis in Appendix B which also discusses a method to tune online recommendations for rationally inattentive user groups to maximize their expected utility.

Context and other Applications

Since Bayesian revealed preferences with rational inattention constitute the main methodology of this paper, below we give some additional insight.

The rationally inattentive model of the decision-making Bayesian agent belongs to a larger class of decision theories based on stochastic information sampling. Models in this class include: (1) *Rational models* (Simon (1955); March (1978)) where information acquisition costs are ignored and agents have time invariant preferences over their actions. (2) *Bounded rationality models* (Caplin and Dean (2015); Sims (2003)) where the Bayesian agent incurs a cost for accurate perception and generates noisy samples conditional on the state (either one-shot or sequentially over time) to take an action for maximizing net expected utility. (3) *Evidence accumulation models* (Krajbich et al. (2010); Ratcliff and Smith (2004)) where consumers' attention is modeled by drift-diffusion models that accumulate evidence based on whether they are fixating their gaze on either the product or its price. The decision is taken when any of the alternatives' evidence threshold level is achieved. (4) *Parallel constraint satisfaction models* (Glöckner and Betsch (2008); McClelland and Rumelhart (1989)) which assume that information is screened sequentially to highlight salient alternatives and final choice is made when the decision maker reaches sufficient internal coherence.

In contrast, examples of visual attention models relevant to our YouTube context from existing literature in psychology include: *(1) Top down control of attention* (Yarbus (2013); Hayhoe et al. (2003)) where agents focus on visual aspects relevant to the decision objective *(2) Bottom up control of attention* (Frintrop et al. (2010); Judd et al. (2012)) where the agent implicitly constructs a topographic saliency map of the visual to guide attention selection.

Testing for Bayesian utility maximization and recovering the resulting utility function also has applications in finance and online marketing.

1. *Online Marketing:* Lee et al. (2009) show that using $9-$ending price displays of online commodities affects consumer choices which were also found to be consistent with rational inattention. Helmers et al. (2015) test whether online consumer choices are consistent with limited attention models. They exploit their results to enhance sales of a particular brand by intelligent product recommendation to the online customers. Using a Bayesian framework and Shannon entropy, Martin (2017) shows how rational inattention of online buyers can be exploited for strategic pricing of goods by online sellers.
2. *Finance:* Shiller (2000) provides evidence that investors are boundedly rational with limited information-processing capability. Huang and Liu (2007) show that costly information acquisition of economic news by Bayesian investors results in their portfolio management being myopic with respect to parameters like news frequency and accuracy. They also char-

acterize desired news frequencies for investors depending on whether they are risk-averse or risk-seeking. This characterization can be used by online financial services companies to selectively advertise economic news to their clients on their online platforms.

Organization

Sec. 2 introduces the problem formulation and connection to constrained contextual bandits. Sec. 3 discusses a deep embedded clustering algorithm for associating the observed agent's action to specific frames. In Sec. 4 and 5, the decision test for rational inattention with behavioral economics constraints stemming from Rényi mutual information are provided. The decision tests are constructive: they provide estimates of the utility function, information acquisition cost, and attention function. Sec. 6 applies the methods to a massive YouTube dataset to characterize the commenting behavior of users. Appendix A contains the proofs for our main theoretical results Theorem 2 and Theorem 4. Appendix B provides Bernstein based finite sample performance bounds. Appendix C summarizes the implementation details of the deep classifier. Appendix D provides additional details about the number of users comprising YouTube user groups. Finally, Appendix E deals with estimating the agent's attention and choice functions.

2. Problem Formulation and Rational Inattention

We first describe the problem formulation from the point of view of the rationally inattentive Bayesian agents; and then from the point of view of the observer (data analyst) that views the dataset generated by the agents. Despite our abstract formulation, the reader should keep in mind the YouTube context outlined above, namely that the rationally inattentive Bayesian agents are groups of YouTube viewers (this is made precise and discussed in detail in Sec. 6), while the observer (data analyst) analyzes the data to test if YouTube viewer groups are rationally inattentive in their commenting behavior.

VIEWPOINT 1. RATIONALLY INATTENTIVE BAYESIAN AGENT

Assume the agent knows the finite state space $\mathcal{X}$ and finite action space $\mathcal{A}$. The agent's prior beliefs of the possible states are given by the prior probability distribution $\mu(x)$, $x \in \mathcal{X}$. The *attention function* $\alpha(s|x)$ of the agent provides a distribution over the signals $s \in \mathcal{S}(\alpha)$ when the state is x. The set of possible signals $\mathcal{S}(\alpha)$ for a given attention function α is finite. The attention function encodes all the information (signals, private information, and measurement mechanism) available to the agent to compute the posterior state distribution. Given the prior $\mu(x)$, and attention function $\alpha(s|x)$, the Bayesian agent computes the posterior distribution as

$$p(x|s) = \frac{\mu(x)\alpha(s|x)}{\sum\limits_{y\in\mathcal{X}} \mu(y)\alpha(s|y)}. \tag{1}$$

The agent has utility function $u(x, a)$ over the states $x \in \mathcal{X}$ and actions $a \in \mathcal{A}$. To give a brief idea of how this abstract model relates to our YouTube dataset, consider a Bayesian agent that characterizes the average viewer/user engagement in a particular group of videos. Each video in this group constitutes an independent trial of the agent. The state corresponds to whether the video is popular with high viewcount or not. The agent does not know the true state at the time of interacting with the video (commenting on, liking or disliking the video). It records an estimate of the state via

the visual cues from the video thumbnail image and description text. This constitutes the agent's signal s.

Definition 1 *An agent satisfies attention rationality if there exists an information acquisition cost $C(\mu, \alpha) \in \mathbb{R}^+$ such that for the selected attention function α, it selects actions $a \in \mathcal{A}$ that satisfy the following conditions:*

i) Expected Utility Maximization:

$$a^* \in \underset{a\in\mathcal{A}}{\operatorname{argmax}}\, \mathbb{E}\{u(x,a)|s\} = \underset{a\in\mathcal{A}}{\operatorname{argmax}} \left\{ \sum_{x\in\mathcal{X}} p(x|s)u(x,a) \right\} \quad \forall p(x|s) \in \mathcal{S}(\alpha) \tag{2}$$

ii) Attention Selection Rationality:

$$\alpha^*(s|x) \in \underset{\alpha\in\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{ \mathbb{E}_{s\in\mathcal{S}(\alpha)}\{\max_{a\in\mathcal{A}}[\sum_{x\in\mathcal{X}} p(x|s)u(x,a)]\} - C(\mu,\alpha) \right\} \tag{3}$$

where $C(\mu, \alpha)$ is the information acquisition cost of attention function α for the prior distribution μ. Also, $\boldsymbol{\alpha}$ denotes the $\mathcal{S}(\alpha) - 1$ dimensional unit simplex of probability mass functions.

Eq. (2) states that the agent selects actions that are consistent with Bayesian utility maximization. In (3), the probability mass function $\alpha^*(s|x)$ is the optimal attention function selected by the agent to maximize the expected utility. In relation to YouTube, the attention function models the attention span of the YouTube viewer when viewing a video before deciding to take a commenting action. The utility can be understood as the agent/commenter's reputation that depends on the popularity of the video, the agent's commenting action a, that is, whether the video gets a high or low comment count, and the average sentiment of the comments (positive, neutral or negative).

In Sec. 5, we will discuss the special case where the information acquisition cost $C(\mu, \alpha)$ is modelled as the Rényi mutual information between the prior μ and the attention function α (see 17). Eq. (3) can then be viewed as a information acquisition cost constrained Bayesian utility maximization problem

$$\alpha^*(s|x) \in \underset{\alpha\in\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{ \mathbb{E}_{s\in\mathcal{S}(\alpha)}\{\max_{a\in\mathcal{A}}[\sum_{x\in\mathcal{X}} p(x|s)u(x,a)]\} \right\}$$
$$\text{s.t. } C(\mu,\alpha) \leq K, \quad K \in \mathbb{R}^+$$

where the Lagrange multiplier of the constraint is set to 1.

VIEWPOINT 2. INVERSE REINFORCEMENT LEARNING: OBSERVER'S MODEL AND DEEP CLUSTERING OF FRAMES

In inverse reinforcement learning, an observer (analyst) seeks to estimate the utility function of a decision maker by observing its actions in response to a stimulus input (Ng et al. (2000)[2]). The

2. Ng et al. (2000) can be viewed as a special case of the NIAS condition (13) defined in Sec. 4. Specifically they consider an infinite horizon discounted cost Markov decision process. Then given the stationary policy and transition probabilities, it is straightforward to construct a set of inequalities that the cost function satisfies. Indeed inverse optimal control dates back to Kalman (1964). The formulation in Sec. 4 (and in Caplin and Dean (2015)) is more general since the framework involves a Bayesian agent (partially observed system where the analyst does not have access to the observations or observation probabilities of the agent) with the added complexity of rational inattention constraints. Despite this added generality, the NIAS (10) and NIAC (11) conditions (Theorem 2) are necessary and sufficient for detecting a constrained Bayesian utility maximizer.

revealed preference framework we consider is more general: by observing the actions of the agent, the observer (analyst) first aims to determine if the agent is rationally inattentive, and if so, estimate the agent's utility function and information acquisition cost. The *action selection policy* of the agent $(\pi(a|x,f))$ is the conditional probability of choosing an action $a \in \mathcal{A}$ given state $x \in \mathcal{X}$ and is defined as:

$$\pi(a|x,f) = \sum_{s \in \mathcal{S}(\alpha)} \eta(a|s)\alpha(s|x), \tag{4}$$

where $\eta(a|s)$ is called the *choice function* i.e. the probability of choosing an action given the signal $s \in \mathcal{S}(\alpha)$. Note that in (4), the probability of choosing an action depends only on the signal realization $s \in \mathcal{S}(\alpha)$ and not only the true state x for a rationally inattentive agent. Thus, the choice function $\eta(a|s)$ in (4) ensures the data is matched i.e. unobserved attention function $\alpha(s|x)$ is consistent with the observed action selection policy. The observer (analyst) has access to the dataset of states x_t and actions a_t chosen by the agent for videos indexed by $t = \{1, \dots, T\}$, where $T = 140,000$ videos (as discussed in Sec. 6):

$$\mathcal{D} = \{(x_t, f_t, a_t)\}_{t=1}^{T}. \tag{5}$$

In (5), the parameter f_t represents all the framing information immediately apparent to the agent. Typically, framing information f_t includes images, video, text, and data. In our YouTube example, f_t maps the title and thumbnail of a video to an integer representing a unique frame. Qualitatively, different values of f_t determine different action policies by the agent for a given title and thumbnail. An important issue when applying rational inattention theory is accounting for the agent's framing effects that impact the agent's behavior. To account for framing effects, we assume there are $\{0, 1, \dots, N\}$ possible frames. In Sec. 3 a deep embedded clustering method is used to construct f_t given the title and thumbnail of the YouTube video indexed by $t = \{1, 2 \dots T\}$, where $T = 140,000$ videos (as detailed in Sec. 6).

Given the set of frames, rational inattention theory aims to determine if the dataset $\mathcal{D}$ is consistent with Definition 1. To test for rational inattention we require estimates of the (possibly randomized) action selection policy $\pi(a|x,f)$ (4) and prior beliefs $\mu(x)$ of the agent. Using dataset $\mathcal{D}$, the maximum likelihood estimates of $\pi(a|x,f)$ and $\mu(x)$ are

$$\hat{\pi}(a|x,f) = \frac{\sum_{t=1}^{T} \mathbf{1}\{x_t = x, a_t = a, f_t = f\}}{\sum_{t=1}^{T} \mathbf{1}\{x_t = x, f_t = f\}}, \quad \hat{\mu}(x) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{x_t = x\}, \tag{6}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. Given the maximum likelihood estimates (6), Sec. 4 provides a decision test for rational inattention; Sec. 5 specifies a test for utility maximization with Rényi mutual information based rational inattention cost of information acquisition. For agents that satisfy the rational inattention test in Theorem 2, methods to recover their utility function $u(x,a,f)$, attention function $\alpha(s|x)$, posterior distribution $s(x)$, and information acquisition cost $C(\mu,\alpha)$ are also provided in Sec. 4.

In Appendix B, we present a finite sample analysis of the agent's action selection policy.

Constrained Contextual Bandits

The rationally inattentive Bayesian model of the agent described in Viewpoint 1 above can be interpreted as a constrained contextual bandit (Badanidiyuru et al. (2014)). In contextual bandits,

the agent receives a reward drawn from a distribution depending on the action and a context vector. In addition, there is an associated cost for each action choice. The agent seeks to maximize its cumulative expected reward less the cost incurred for its action choices. In our YouTube case, the arm pulled by the agent is equivalent to the attention function $\alpha(s|x)$ chosen by the agent (i.e. we assume a continuum of arms), the cost of pulling an arm becomes the behavioral economics-based rational inattention cost $C(\mu, \alpha)$ and the context corresponds to decision problems defined in Sec. 6. The expected reward for choosing action $\alpha(s|x)$ given context k (decision problem) is the objective function in (3) which is maximized by the agent. Having a continuum of arms is more general than a finite arm bandit problem. Indeed, the conditions in Theorem 2 still remain necessary and sufficient for utility maximization for a discrete set of admissible attention functions (arms).

Viewpoint 2 corresponds to inverse reinforcement learning of a contextual bandit with rational inattention constraints. As the observer (analyst), our aim is to decide if the agent is a constrained Bayesian utility maximizer and if so, reconstruct the utility function of the agent by observing its actions. Thus, as observers (analysts), we do inverse reinforcement learning of a constrained contextual bandits problem. To the best of our knowledge, inverse reinforcement learning for constrained contextual bandits has not been addressed in the literature.

3. Constructing Preference and Policy Invariant Frames via Deep Embedded Clustering

Recall from Sec. 1 that in behavioral economics, "framing" describes information an agent has when making a decision. Framing information can dramatically impact the agent's action selection policies in YouTube. Specifically, the title and thumbnail have a significant impact on the agent's commenting behavior as reported in (Hoiles et al. (2017)). Such framing effects must be accounted for to minimize the probability of incorrectly rejecting a rationally inattentive agent. In Sec. 6, we interpret the YouTube dataset partitioned into groups of videos based on the framing information as individual agents to test if utility maximization holds in these frames.

To account for these framing effects a deep embedding method is provided that learns the policy invariant frames of the agent. Specifically, a mapping of f_t to $n_t \in \{1, \ldots, N\}$ is constructed where for each $n \in \{1, \ldots, N\}$ the behavior of the agent is invariant. In the YouTube social network the framing information available to the agent is comprised of the title and thumbnail of each video. Given that agents are preference invariant to minor variations in the title and thumbnail, it is possible to map the features f_t to one of $\{1, \ldots, N\}$ discrete frames learned using deep embedding. The deep embedded clustering method achieves two objectives: (i) Converts raw image and text data from YouTube videos to a 2-dimensional vector in the Euclidean space and (ii) Partitions the 140000 videos into 4 non-overlapping clusters.

The deep embedding method uses natural language processing and image processing tools in an autoencoder (see Appendix C) to construct the latent representation z_t of f_t, and includes a clustering layer to simultaneously learn how to associate each f_t to one of $\{1, \ldots, N\}$ discrete frames. A schematic of the clustering method is illustrated in Fig. 1.

The autoencoder comprises two deep neural networks, the first is the encoder that maps the input f_t to the latent space representation z_t, and the second is the decoder that map the latent space representation z_t to the input f_t where $\hat{f}_t \approx f_t$. To force the encoder to learn robust latent representations, the autoencoder is trained using corrupted versions of the input with the noise artifically fed from the "Noise" block in Fig. 1. Such an autoencoder is known as a denoising

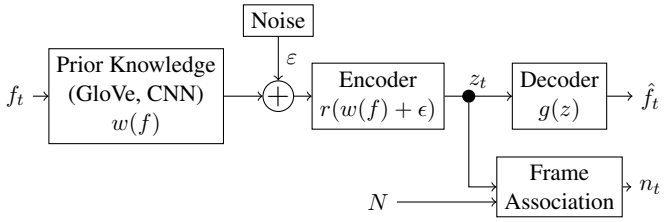

Figure 1: Schematic of the deep embedded clustering method to map the framing information f_t to the discrete frame $\{1, \ldots, N\}$. The parameter $w(f)$ contains all prior knowledge of the input framing information, ε is a Gaussian white noise term, $r(\tilde{f})$ is the encoder, z_t is the latent space representation of f_t, $g(z)$ represents the decoder, and $\hat{f}_t$ is the output of the autoencoder.

autoencoder (Vincent et al. (2008); Bengio (2009)). The denoising autoencoder encodes the input into the latent space representation as a $200-$dimensional vector as $z_t = r(w(f_t) + \epsilon)$, and attempts to remove the effect of this corruption process stochastically applied to the input of the autoencoder. Removing effects of the corruption process is performed by learning the statistical dependencies between the inputs. A detailed description of the denoising autoencoder architecture is in Appendix C with focus on the title and thumbnail of YouTube videos.

Though the latent space representation of the input has been used extensively for clustering, such methods are not guaranteed to preserve any intrinsic local structure of the framing data f_t. To ensure the autoencoder both minimizes the reconstruction error and maximizes the intrinsic local structure of the data, a clustering loss is used. The loss of the deep embedded clustering method (Fig. 1) is:

$$L = \left(\sum_{t=1}^{T} ||f - g(r(w(f) + \varepsilon))||_2^2 \right) /T + \mathrm{KL}(P||Q), \tag{7}$$

where $g(\cdot)$ is the decoding function of the denoising autoencoder and $T = 140000$. The first term in (7) is the reconstruction error, and $\mathrm{KL}(P||Q)$ denotes the Kullback-Leibler (KL) divergence of the discrete probability distributions P and Q. Here Q (with elements $q_{t,n}$) is the prior probability distribution of cluster association between the latent variables z_t and the associated frames n_t. If we assume each cluster is generated from a Gaussian normal distribution with mean Ψ_n, then the normalized probability of association of each z_t with every cluster head Ψ_n is given by:

$$q_{t,n} = \frac{1}{T} \left[\frac{(1 + ||z_t - \Psi_n||^2)^{-1})}{\sum_{n=1}^{N} (1 + ||z_t - \Psi_n||^2)^{-1}} \right] \quad \forall n \in \{1, \ldots, N\}. \tag{8}$$

Note that the above definition ensures $\sum_{t=1}^{T} \sum_{n=1}^{N} q_{t,n} = 1$. Given Q, to avoid degenerate clustering solutions which allocate most of the frames to a few clusters or assign a cluster to a sample outlier, the target distribution P is designed with elements $p_{t,n}$ defined as:

$$p_{t,n} = \frac{1}{T} \left[\frac{q_{t,n}^2 / \sum_{t=1}^{T} q_{t,n}}{\sum_{n=1}^{N} (q_{t,n}^2 / \sum_{t=1}^{T} q_{t,n})} \right], \quad P(z_t = n) = \sum_{t=1}^{T} q_{t,n}, \quad t = 1, \ldots, T \tag{9}$$

Note that $P(z_t = n)$ is the empirical frequency that z_t belongs to cluster n.

From (8) and (9), if all the data-points are associated with a specific cluster this will increase the loss (7). Additionally, if the cluster is associated with several data points with low-confidence, this will also increase the loss (7). Minimizing the loss (7) can be interpreted as a form of self-training as P depends on Q. Specifically, in self-training we take an initial classifier and an unlabeled dataset, then label the dataset with the classifier in order to train on its own high confidence predictions. This ensures that the latent clusters are constructed to avoid outliers.

The deep embedding method that maps f_t to $n_t \in \{1, \ldots, N\}$ is formalized in Algorithm 1; see Algorithm 2 in Appendix C for the detailed version. The pre-training step is used to initialize the encoder and decoder parameters prior to performing any clustering. This is a critical step as the initial latent space representation of $\{f_t\}_{t=1}^T$ is used to select the approximate locations of the N latent space cluster centers Ψ^o. Given the pre-trained denoising autoencoder weights, we use the Lloyd heuristic algorithm to select the locations of the N latent space cluster centers Ψ^o. Given the cluster centers, the deep clustering method is applied to minimize the loss (7) by simultaneously adjusting the cluster associations and autoencoder weights. Note that in Algorithm 2, since the distribution P (9) depends on the weights of the encoder, we update P after ζ iterations. This reduces the probability of instability associated with cycling between adjusting weights and cluster associations. The final result of Algorithm 2 is achieved when the change in cluster associations is below a threshold δ.

Algorithm 1 Deep Embedded Clustering for Framing Association

Require: Set of framing information $\{f_t\}_{t=1}^T$, number of unique frames N, stopping threshold $\delta \in (0, 1)$, confidence threshold $\delta_c \in (0, 1)$, and updating interval ζ.

PRE-TRAIN

Pre-train the denoising autoencoder without any frame association.

INITIALIZE

Initialize the N cluster centers Ψ^o using k-means clustering in the latent space and set $\varepsilon = 0$.

DEEP CLUSTERING

Train the deep clustering autoencoder and frame association layers (refer to Appendix C).

return Invariant frames $n_t \forall t \in \{1, \ldots, T\}$ such that $\max_n\{q_{t,n}\} > \delta_c$.

Given the preference and policy invariant frames $\{n_t\}_{t=1}^T$, we substitute $n_t \to f_t$ in $\mathcal{D}$ (5). Within each frame $f \in \{1, 2 \ldots N\}$, we further divide the group of videos into K sub-partitions or *decision problems*. The procedure for further partitioning each frame is explained in detail in Sec. 6.1. For each unique frame $f \in \{1, 2, \ldots N\}$, we assume the videos with frame index f to be generated from a unique preference ordering; the unique frames are henceforth interchangeably referred to as *invariant* frames. Using $\mathcal{D}$ with the invariant frames and their corresponding decision problems, Sec. 4 and Sec. 5 illustrate how to detect if the agent is rationally inattentive for different information acquisition cost constraints, and how to recover the utility functions.

Remark: The deep learning based denoising autoencoder extracts the most important features of the video thumbnail and description that cannot be learned using heuristic methods. This also offers a greater level of sophistication for the dimension reduction of the dataset compared to conventional dimension reduction methods liks PCA (see related works like Mikolov et al. (2010)).

4. Decision Test for Rational Inattention; Estimating Utility and Information Acquisition Costs

Here we construct a decision test for rational inattention (Definition 1) that involves estimating the utility as well as the information acquisition costs. The resulting preference-based inverse reinforcement learning algorithm uses the observed stochastic choice dataset $\mathcal{D}$ (5) and invariant frames $f \in \{1, 2 \dots M\}$. Recall that the group of YouTube videos in each frame is further subdivided into K decision problems and their associated utility functions $\{u_k(x,a)\}_{k=1}^K$ as defined in (2). Theorem 2 below is our main result and is a slight generalization of Caplin and Dean (2015) and Caplin and Martin (2015). The proof is in Appendix A.

Theorem 2 *The dataset $\mathcal{D}$ (5) satisfies rational inattention (Definition 1) iff the action selection policy $\pi_k(a|x,f)$ defined in (4) and utility function $u_k(x,a,f)$ (2) for decision problems (d.p.) $k \in \{1, 2 \dots K\}$ for every frame $f \in \{1, 2 \dots N\}$ satisfy the following two conditions:*

(1) ***NIAS:*** $\sum_{x \in \mathcal{X}} p_k(x|a,f)[u_k(x,a,f) - u_k(x,b,f)] \geq 0 \quad \forall a, b \in \mathcal{A}_k$

$$p_k(x|a,f) = \frac{\mu(x)\pi_k(a|x,f)}{\sum_{y \in \mathcal{X}} \mu(y)\pi_k(a|y,f)} \tag{10}$$

(2) ***NIAC:*** $\sum_{k=k_1}^{k_L} G_{k,k,f} - G_{k+1,k,f} \geq 0$ *for any sequence of d.p.* $k_1, k_2 \dots k_L (L \leq K)$ (11)

$$G_{k,w,f} = \sum_{s \in \mathcal{S}(\alpha_k)} \sum_{x \in \mathcal{X}} \mu(x)\pi_k(a|x,f) \max_{b \in \mathcal{A}_w} \left\{ \sum_{x \in \mathcal{X}} p_k(x|a,f) u_w(x,b,f) \right\} \ (k_{L+1} = k_1)$$

Theorem 2 is proved in Appendix A. It specifies necessary and sufficient conditions for an agent to be a Bayesian utility maximizer (Caplin and Dean (2015)), namely, *No Improving Attention Switches* (NIAS) and *No Improving Attention Cycles* (NIAC).

Theorem 2 is best viewed from the point of view of an observer (analyst) performing inverse reinforcement learning on a Bayesian agent. If (10) and (11) have no feasible solutions, then the agent does not satisfy rational inattention based Bayesian utility maximization. According to Theorem 2, the analyst only needs to know the action selection policy $\pi_k(x|a,f)$ defined in (4); this can be estimated from the dataset as described in (6). Note that the analyst does not require knowledge of the agent's attention function $\alpha_k(s|x,f)$ defined in (3) even though this attention function together with the utility $u_k(\cdot)$ determines the actions of the k^{th} agent.

A few words about NIAS and NIAC. The NIAS condition implies that the agent's actions are optimal under posterior beliefs. The NIAC condition implies that the sum of expected utilities (3) over decision problems cannot be increased by reassigning attention selection policies. For readers familiar with revealed preference theory, the NIAC condition is analogous to the GARP conditions in Afriat's theorem (Varian (2012); Diewert (2012)) for testing utility maximization behavior. $G_{k,w}$ in (11) gives the expected utility of using attention function $\alpha_k(s|x,f)$. Additionally, one can construct the total expected utility of an agent given by

$$V(\boldsymbol{\pi}(a|x,f)) = \sum_{k=1}^{K} \sum_{x \in \mathcal{X}} \sum_{a \in \mathcal{A}_k} \pi_k(a|x,f)\mu(x)u(a,x,f). \tag{12}$$

As an aside, in Appendix B we illustrate how the observer (analyst) performing inverse reinforcement learning can maximize $V(\boldsymbol{\pi})$ w.r.t policies $\boldsymbol{\pi}$ so as to provide optimal behavioral recommendations to agents.

From a practical point of view, (10), (11) in Theorem 2 can be implemented by solving the following equivalent convex feasibility problem:

Corollary 3 *The NIAS and NIAC conditions in Theorem 2 can be equivalently expressed as the following convex feasibility problem (convex in the utility functions u_k):*
Find $u_k(x,a,f) \in [0,1],\ \forall a,b \in \mathcal{A}_k,\ \forall k \in \{1,2\dots K\}, \forall f \in \{1,2\dots N\}$ such that

$$\textbf{\textit{NIAS}}: \quad \sum_{x\in\mathcal{X}} p_k(x|a,f)[u_k(x,b,f) - u_k(x,a,f)] \leq 0$$

$$\textbf{\textit{NIAC}}: \quad \sum_{k=k_1}^{k_L}\Bigg(\sum_{a\in\mathcal{A}_{k+1}}\Bigg(\sum_{x\in\mathcal{X}}\mu(x)\pi_{k+1}(a|x,f)\Bigg)\max_{a\in\mathcal{A}_k}\sum_{x\in\mathcal{X}}p_{k+1}(x|a)u_k(x,a,f) - \sum_{a\in\mathcal{A}_k}\sum_{x\in\mathcal{X}}\mu(x)\pi_k(a|x,f)u_k(x,a,f)\Bigg) \leq 0,\ \forall k_{1:L} \in \{1,\dots K\} \tag{13}$$

where $k_{L+1} = k_1$ and $L \leq K$.

Corollary 3 provides a constructive feasibility test for the analyst to determine if YouTube user groups in dataset $\mathcal{D}$ satisfy rational inattention based Bayesian utility maximization. It also provides a set-valued estimate of the user groups' utility function $u_k(x,a,f)$ that rationalizes the dataset $\mathcal{D}$. To justify our findings on the YouTube dataset, in Sec. 6 we perform robustness tests for every segment (groups of videos) in the YouTube dataset $\mathcal{D}$ (see Sec. 6.2 and Sec. 6.3 for partitioning of the dataset into segments) to see how close they are to satisfying utility maximization behavior.

Algorithms for convex feasibility Boyd and Vandenberghe (2004) can be used to check feasibility of (13) and construct a feasible solution.

Construction of Information Acquisition Cost: Given the utility function $\{u_k(x,a,f)\}_{k=1}^K$ that satisfy the constraints (12) and (13), the observer (analyst) can then construct an estimate of the associated cost of information acquisition $C(\mu,\alpha_k)$ of each attention function α_k. Specifically, the estimate of $C(\mu,\alpha_k)$ can be computed by solving a second convex feasibility problem:

$$\begin{aligned} &G_{k,k,f} - G_{w,k,f} \geq C_f(\mu,\alpha_k) - C_f(\mu,\alpha_w) \\ &C_f(\mu,\alpha_k) \geq 0 \quad \forall w,k \in \{1,\dots,K\}, \end{aligned} \tag{14}$$

where $G_{k,w,f}(\cdot)$ is the expected utility computed in Corollary 3.
Remark: The set of all feasible utility functions and information acquisition costs $\{u_k(x,a,f)$, $C(\mu,\alpha_k), k \in \{1,2,\dots K\}, f \in \{1,2,\dots N\}\}$ w.r.t inequalities (13), (14) is closed under scalar multiplication by a positive real number. These estimates are thus 'ordinal' estimates since any positive scalar multiple of the estimates explains the dataset $\mathcal{D}$ equally well.

Recall that if a solution to (12) exists, then a solution to (14) is guaranteed to exist from Theorem 1 and (3). Also, if the cost of a particular attention function is zero, then absolute bounds can be placed on the information acquisition cost of each attention function. For example if $C(\mu,\alpha_w) = 0$, then the cost $C(\mu,\alpha_k) \in [G_{k,w} - G_{w,w}, G_{k,k} - G_{w,k}]$. The estimated cost function satisfies weak monotonicity in information–that is, if the attention function provides more information then it incurs a higher information acquisition cost.

5. Rényi Entropy Information Acquisition Cost for Rational Inattention

In this section we impose a behavioral economics structure to the information acquisition cost $C(\mu, \alpha)$ in (3) which defines the attention function of a rationally inattentive agent. Sims' pioneering work (Sims (2010)) uses Shannon mutual information between the prior distribution and attention function (α) to define information acquisition cost in such agents whereas here the more general Rényi mutual information is considered. The Rényi mutual information between the prior $\mu(x)$ of the state and the selected attention function $\alpha_k(s|x)$ (k denotes the k^{th} decision problem (10)) is

$$I_\beta(\mu, \alpha_k) = \begin{cases} \frac{1}{\beta - 1} \ln \left(\sum_{x \in \mathcal{X}} \sum_{a \in \mathcal{A}} \frac{p^\beta(x,a)}{\mu^{\beta-1}(x) p^{\beta-1}(a)} \right) & \beta \in (0,1) \cup (1, \infty) \\ I(\mu, \alpha_k) \quad \beta = 1 & \\ - \ln \left(\sum_{x \in \mathcal{X}} \sum_{a \in \mathcal{A}} \mu(x) p(a) \mathbb{1}\{p(x,a) > 0\} \right) & \beta = 0 \end{cases} \tag{15}$$

where $\beta \in [0, \infty)$ is the Rényi order. Note that we replace the signal "s" with the action "a" since we assume a one-to-one mapping between the two (parsimonious representation). The mutual information defined in (15) stems from the Rényi entropy defined as

$$H_\beta(X) = \frac{1}{1 - \beta} \log \big(\sum_{i=1}^{n} p_i^\beta \big). \tag{16}$$

The Rényi entropy is useful for measuring the information acquisition cost since the parameter β allows one to adjust the sensitivity of the cost to the shape of $\mu(x)$ and $\alpha_k(s|x)$. Indeed, for suitable choice of β the Rényi entropy includes the Hartley entropy, Shannon entropy, collision entropy and minimum entropy as special cases.

An important feature of (15) is that for $\beta \in (0, 1]$ the mutual information constraint is convex in $p(x, a)$ and $\mu(x)p(a)$. Also for $\beta > 1$, the information constraint is convex in $\mu(x)p(a)$ and quasi-convex in $p(x, a)$ (Van Erven and Harremos (2014); Ho and Verdú (2015); Xu and Erdogmuns (2010)). In terms of (2), the Rényi mutual information cost constrained decision problem is

$$\begin{aligned} p_k^*(x,a) \in \underset{p(x,a)}{\operatorname{argmax}} & \Big\{ \sum_{a \in \mathcal{A}_k} \sum_{x \in \mathcal{X}} p(x,a) u(x,a) \Big\} \\ \text{s.t.} \quad & \mu(x) = \sum_{a \in \mathcal{A}_k} p(x,a) \quad \forall x \in \mathcal{X} \\ & I_\beta(\mu, \alpha_k) \leq \kappa_{\text{max,k}}, \quad p(x,a) \geq 0 \quad \forall x \in \mathcal{X}, a \in \mathcal{A}_k. \end{aligned} \tag{17}$$

In (17), κ_{max} represents the maximum "effort" the agent is willing to invest to estimate the state $x \in \mathcal{X}$ prior to taking the action $a \in \mathcal{A}_k$ in decision problem $k \in \{1, \ldots, K\}$. The constraint in (17) constitutes the behavioral economics based constraint.

In addition to satisfying the NIAC and NIAS conditions in Theorem 2, the information acquisition cost $C(\mu, \alpha)$ in Definition 1 is now defined as:

$$C(\mu, \alpha) = \lambda I_\beta(\mu, \alpha) + \gamma, \tag{18}$$

where $\lambda \in \mathbb{R}^+, \gamma \in \mathbb{R}, \beta \in [0,1]$ and $I_\beta(\mu, \alpha_k)$ is as defined in (15). Eq. (18) denotes the Rényi Mutual Information cost and Shannon Mutual Information cost for $\beta \in (0,1)$ and $\beta = 1$ respectively. In (18), the objective function is linear and the constraint set is convex in the argument $p(x,a)$ for $\beta \in [0,1]$, hence necessary and sufficient conditions for the agent to satisfy rational inattention with the Rényi/Shannon mutual information cost (18) can be constructed using the Karush-Kuhn-Tucker (KKT) conditions. Formally:

Theorem 4 *Consider a Bayesian agent that satisfies the conditions (10) and (11) in Theorem 2. This agent satisfies utility maximization for*

(i) ***Rényi mutual information cost*** *(18) iff there exist constants $\lambda_{1,k} > 0$ and $\lambda_{2,k}$ that satisfy the linear constraints*

$$u_k(x,a) = \lambda_{1,k} \frac{\beta \eta^{\beta-1}(x,a) - (\beta-1)\mathbb{E}_x\left[\eta^{\beta-1}(x,a)\right]}{(\beta-1)(\mathbb{E}\left[\eta^{\beta-1}(x,a)\right])p^{\beta-1}(a)} - \lambda_{2,k},$$
$$\frac{1}{\beta-1}\ln\left(\mathbb{E}[\eta_k^{\beta-1}(x,a)]\right) = \kappa_{max}, \qquad \eta_k(x,a) = \frac{p_k(x|a)}{\mu(x)}, \tag{19}$$

for all $k \in \{1,2\ldots K\}, x \in \mathcal{X}$, $a \in \mathcal{A}$.

(ii) ***Shannon mutual information cost*** *(18) iff there exist constants $\lambda_{1,k} > 0$ and $\lambda_{2,k}$ that satisfy the linear constraints*

$$u_k(x,a) = \lambda_{1,k}\ln(p_k(x|a)) - \lambda_{2,k}, \quad \mathbb{E}\left[\ln\frac{p_k(x,a)}{\mu(x)p_k(a)}\right] = \kappa_{max}, \tag{20}$$

for all $k \in \{1,2\ldots K\}, x \in \mathcal{X}$, $a \in \mathcal{A}$.

The proof is in Appendix A.2. In Theorem 4, $\lambda_{1,k} > 0$ and $\lambda_{2,k}$ are Lagrange multipliers pertaining to the inequality and equality constraint in (17). Combining the linear equality constraints (19), (20) with the convex feasibility problem in Corollary 3 yields a test for the Rényi/ Shannon mutual information cost constrained optimization problem and provides estimates of the associated utility function of the agent. Thus we have constructed a preference based inverse reinforcement learning algorithm for the utility and information acquisition cost of a Bayesian agent with a Rényi and Shannon mutual information cost.

6. Rationally Inattentive Inverse Reinforcement Learning in YouTube Engagement

This section provides empirical evidence that the commenting behavior of YouTube users is consistent with Bayesian utility maximization with rational inattention. We consider a massive YouTube dataset comprising approximately 140000 videos across $25,000$ channels and over 9 millions users from April 2007 to May 2015.

YouTube is an interesting example of a social network since the interaction between users includes video content. Users interact on YouTube channels by posting comments and rating videos. In this paper, we analyze overall user engagement in YouTube videos. By inverse reinforcement learning in YouTube, we mean determining the existence and construction of utility functions in various user groups that rationalizes the commenting behavior in the YouTube dataset $\mathcal{D}$. Extensive empirical studies (Khan (2017a); Kwon and Gruzd (2017); Alhabash et al. (2015); Hoiles et al.

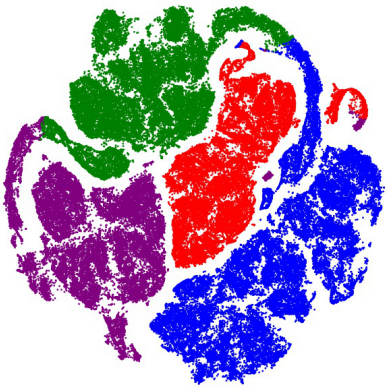

Figure 2: t-SNE visualization (Van der Maaten and Hinton (2008)) of the latent space representation of title and thumbnail of YouTube videos constructed using Algorithm 2 for the YouTube videos contained in the dataset $\mathcal{D}$. As seen, the t-SNE representation with 4 frames (each indicated by a different color) indicates that the videos are sufficiently separated in the latent space.

(2015, 2017); Aprem and Krishnamurthy (2017)) show that the comments and ratings from users are influenced by the thumbnail, title, category, and perceived popularity of each video.

Before proceeding with details of our analysis, we briefly summarize our main conclusions:

1. Rationally inattentive commenting behavior (Definition 1) exists in the majority of YouTube user groups.
2. Our decision test is robust- The segments that do not pass the rational inattention test (Theorem 2) are close to satisfying the inequalities in (13).
3. The utility function constructed from the rational inattention test (Corollary 3) can be used to accurately predict commenting behavior (with 83% accuracy).

The results presented in Sec. 6.2 and Sec. 6.3 of this paper can be reproduced using the code and datasets that we have uploaded to a public GitHub repository: `https://github.com/KunalP117/YouTube_project`.

6.1 YouTube Dataset and Model Parameters

The massive YouTube dataset that we analyze comprises $140,000$ videos across 25,000 channels from April 2007 to May 2015. We index the videos by $t = \{1, 2, \dots T\}$, where $T = 140,000$. The dataset contains the view counts, comment counts, likes, dislikes, thumbnail, title, and category of each video. To relate to our main Theorem 2, we define the following:

1. *Agent*: Group of users (See Appendix D for more details on the number of users per group) interacting with videos in each video segment. Each agent is associated with a decision problem (either a frame or category as defined below).

2. *State (x_t)*: In the YouTube dataset, the state x_t of each video is the viewcount 1 day after the video was published[3]. Specifically, state $x_t = 1$ is high viewcount (more than $10,000$ views) and $x_t = 2$ otherwise.

3. *Action (a_t)*: In the YouTube dataset, the associated action a_t is related to the overall commenting behavior[4] of the agents, which is computed using the comment counts, like count, and dislike count 2 days after the video is published. The possible actions are: $a_t = 1$ denotes low comment count with negative sentiment, $a_t = 2$ denotes low comment count with neutral sentiment, $a_t = 3$ denotes low

3. Alternatively, we could have used the number of subscribers when the video was published as the state. Our extensive data analysis in Hoiles et al. (2017) shows that the viewcount after 1 day and the initial subscriber count are the two highest sensitivity features for estimating the viewcount 14 days after the video is published.

4. By overall commenting behavior in YouTube, we mean both the comment count and the video ratings (likes and dislikes). Another term used in the literature (Khan (2017b)) is "user engagement".

comment count with positive sentiment, $a_t = 4$ denotes high comment count with negative sentiment, $a_t = 5$ denotes high comment count with neutral sentiment, and $a_t = 6$ denotes high comment count with positive sentiment. Here negative sentiment occurs if the difference between the like count and dislike count is less than -25, neutral sentiment occurs if the difference lies between $-25, 25$, and positive sentiment occurs if the difference is greater than 25. A low comment count is said to occur if there are less than 100 comments, otherwise the comment count is defined to be high. If the user searches for a video, or directly goes to the video webpage, then the thumbnail image and viewcount are the first pieces of information that the user sees. The user needs to either click on the thumbnail or scroll down the specific video webpage (both actions affect the video's viewcount) in order to post their comments.

4. *Frame (f_t)*: The frame f_t of a video refers to the segment each video in the YouTube dataset is categorized into by Deep Embedded Clustering based on the video's framing information and is indexed as $1...N$. The framing information for each video is comprised of the video's thumbnail and

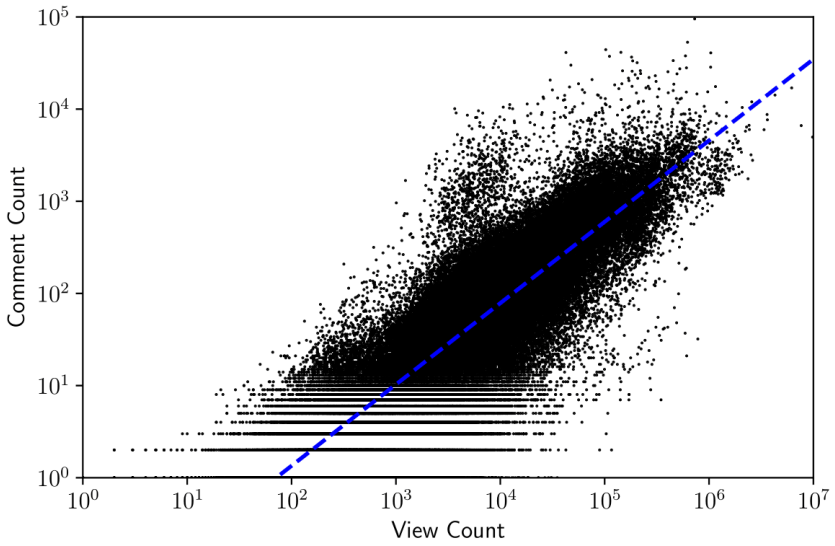

Figure 3: The comment count vs. viewcount follows a power law with power coefficient of 0.88, namely, $c = 0.023 \times v^{0.88}$ in the YouTube dataset $\mathcal{D}$. So, the comment count increases sub-linearly and the rate of change of commenting decreases with increasing viewcount.

title. Specifically, we use a 40×80 pixel color image to represent the thumbnail (which is a resized version of the native 246×138 pixel thumbnails used in YouTube). For the title, we only include the first 8 words of the title in the framing instance f_t (over 90% of the videos have a title of length 8 words or less).

5. *Decision problem (k)*: (i) When data is partitioned by framing (via deep clustering methods of Sec. 3), for each frame, a decision problem corresponds to one of two types the video belongs to - gaming videos or non-gaming videos (Sec. 6.2). (ii) When data is partitioned by the video category (based on Sec. 6.3), the decision problem is specified by the category each video belongs to. Each video contains a category index $k \in \{1, \ldots, 18\}$ representing the specific YouTube category descriptor of the video. Fig. 7 in Appendix D lists each video category with the total number of views. Note that the video categories "Unavailable" or "Removed" are videos flagged by YouTube as being suspected

of violating YouTube's video policies[5].
6. *Information Acquisition Cost*: We consider the following parameterized costs of acquisition of attention function α_k used by the k^{th} Bayesian agent (3): (i) General cost (14), (ii) Rényi mutual information cost (15), (iii) Shannon mutual information cost (15)

To give additional visual insight (although tangential to the paper), Fig. 3 displays the comment count vs. viewcount of $140,000$ videos comprising the dataset $\mathcal{D}$. As shown, a power law yields approximately 2% smaller residual root mean squared error than a linear fit. It can be seen that while the associated comment count increases with view count, the rate of change of commenting decreases with increasing viewcount. Of course, this paper and the data analysis below focus on the different issue of how to determine a Bayesian utility function that rationalizes comments for groups of users given the viewcount.

6.2 YouTube Data Analysis 1: Deep Clustering Approach

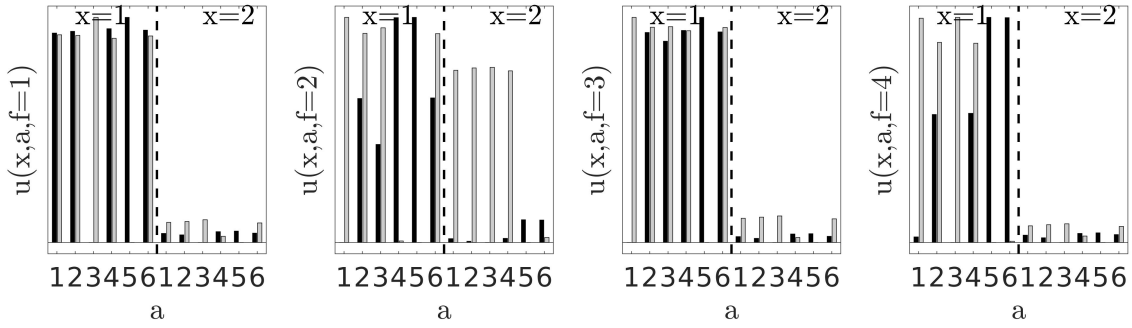

Figure 4: Utility functions $\{u_k(x,a,f)\}_{k=1}$ of the rationally inattentive agents with a general cost for information acquisition for each of the four unique frames are plotted in Fig. 4. The utility is constructed by evaluating the convex feasibility problem (13) for the YouTube dataset $\mathcal{D}$. x represents the state, a the possible actions, f the frame, and the decision-problem k indicates the most popular category (black bars) and the other categories (gray bars). The main conclusions are that a utility function can be constructed that explains the commenting preferences of users, which also indicates that a typical YouTube commenter has a higher preference (utility) to comment on videos with a higher viewcount (popularity metric) for all actions $a \in \{1,2,3,4,5,6\}$. Recall that these actions reflect the comment count and sentiment of comments.

Our first approach for analyzing the YouTube dataset is user centric: it involves constructing frames over which preference ordering for nature of commenting behavior stays invariant using deep embedded clustering via Algorithm 2. Recall that the frame f_t of a video refers to the segment each video in the YouTube dataset is categorized into by Deep Embedded Clustering based on the video's extrinsic framing information. Algorithm 2 maps the high dimensional title and thumbnail space to one of N unique frames. Fig.2 depicts a two stage dimension reduction described as follows: The raw high dimensional vectors comprising video description and thumbnail are first mapped to a latent embedding space (of dimension 200) as part of Algorithm 2. The points in the latent

5. Refer to `https://www.youtube.com/yt/about/policies/#community-guidelines` for details

embedding space are further projected to a 2-D space for better visualization (t-distributed Stochastic Neighbor Embedding (Van der Maaten and Hinton (2008))). Each of the four colors in the figure represents a distinct frame. Choosing $N = 4$ ensures each video is sufficiently isolated to a particular frame; less than 3% of videos are classified ambiguously in terms of frames. The most popular category of videos in the YouTube dataset is "Gaming" which comprises 44% of all the videos. Two decision-problems are considered within each frame of the dataset. The first is $k = 1$ which is associated with all videos that have category "Gaming", while decision problem $k = 2$ results for videos that are not associated with the "Gaming" category.

Data Analysis Results: We apply the rational inattention test (13) from Corollary 3 to each of the preference invariant frames in Fig. 2. Our main conclusion is that the commenting behavior of users in YouTube is consistent with rational inattention for a general cost constraint (14). The utility of the users in each unique frame (Fig. 2) is provided in Fig. 4. It can be inferred from the constructed utility function in Fig. 4 that users prefer to comment on videos that are expected to have a higher popularity compared to videos with lower popularity.

The associated utility however provides no clear preference ordering between the popularity of the video and the associated commenting behavior. This suggests that YouTube user groups in each frame are rationally inattentive with respect to a general cost. If we impose the Rényi Mutual Information cost constraint (18), we find that only the commenting behavior in frame $f = 4$ is rationally inattentive. It was also found that frame 4 satisfies rational inattention for a Shannon Mutual Information cost constraint (18). Thus, the user groups in frame 4 satisfy utility maximization for all information acquisition costs.

6.3 YouTube Data Analysis 2 : Category Dissection Approach

We now describe a second approach for analyzing the YouTube dataset; this approach is data centric and uses YouTube video categories instead of frames. The diversity of videos in YouTube is immense, and it is natural to exploit this diversity for prediction of attributes in YouTube commenting behavior. Hence, we now analyze YouTube data by partitioning groups of videos based on their category. The aim is to determine if a utility function rationalizes each category of YouTube videos. Categories in YouTube (eg. News, Gaming, Music etc.) are numbered from $1 - 18$ (See Fig. 7 for the full listing). The granularity of the analysis and prediction in this section is much finer than the analysis in Sec. 6.2. As discussed in Appendix D, the video categories have mean numbers of users ranging from 149 to 4596 for high viewcount (greater than 10000) videos and 8 to 1801 for low viewcount videos (less than 10000).

In the current formulation, we apply the rational inattention test (Corollary 3) on each distinct pair of video categories (i, j) from the existing category index set. This implies the number of decision problems $K = 2$ in (10). Each decision problem corresponds to the distinct video categories i and j and the variable f in Theorem 2 indexes the video category pair (i, j). The main findings of this section are:

1. Rationally inattentive commenting behavior holds for approximately 56% of the video categories in YouTube.
2. Our decision test for rational inattention is robust to parameterization in information acquisition cost- with a small perturbation (quantified precisely below), every user group can be rationalized by a utility function irrespective of the information acquisition cost.

Information Acquisition Cost	Number of category pairs from $\binom{18}{2}$ pairs that satisfy rational inattention decision test (Corollary 3)	Set of categories
General (14)	45	$\{1, 5, 7, 8, 9, 10, 11, 12, 14, 17\}$
Rényi Mutual Information (18)		
$\beta \in (0, 0.7)$	0	$\{\}$
$\beta \in [0.7, 0.83)$	2	$\{11, 12\}$
$\beta \in [0.83, 1)$	3	$\{11, 12, 17\}$
Shannon Mutual Information (18)	2	$\{11, 17\}$

Table 1: Categories of YouTube videos where the commenting behavior satisfies Bayesian utility maximization behavior

3. We construct utility function for specific video categories in YouTube and use these utility functions to predict (with 83% accuracy) commenting behavior in those categories.

6.3.1 YOUTUBE DATASET ANALYSIS FOR UTILITY MAXIMIZATION WITH GENERAL COST CONSTRAINT

Utility maximization for Rényi/Shannon Mutual Information cost implies utility maximization with the unconstrained general cost. The results are displayed in Table 1 where it is shown that approximately 56% of categories satisfy the general cost. 90% of videos in the YouTube dataset belong to this set of categories. For Rényi Mutual Information cost (18), we see that the number of pairs of video categories satisfying utility maximization increases with β.

6.3.2 ROBUSTNESS OF RATIONALLY INATTENTIVE UTILITY MAXIMIZATION TEST

A natural robustness question is: For the categories of YouTube data that failed the utility maximization test of Theorem 1, how close are these categories to passing the utility maximization test? Similarly, for the categories in YouTube passing the utility maximization test, how far are these categories from failing the test? The purpose of this section is to define three robustness measures: one for the video categories in YouTube data that do not satisfy Theorem 1 and two for the categories in YouTube which do satisfy Theorem 2. The main result below is that for the general cost for information acquisition (14), these robustness measures are large for video category pairs that satisfy the utility maximization test and are relatively small for the category pairs that fail the test, for all 18 video categories in YouTube. This implies that the commenting behavior in video categories that satisfy utility maximization do so by a large margin; and the video categories that fail utility maximization do so by only a small margin. Thus one can conclude that the commenting behavior across *all* video categories in YouTube is approximately rational. Recall from Sec. 6.1 that in our YouTube data analysis, the state space $\mathcal{X} = \{1, 2\}$ and action space $\mathcal{A} = \{1, 2, 3, 4, 5, 6\}$. Let $C = \{(i, j) | i \neq j, i, j \in \{1, 2.., 18\}\}$ denote the set of all possible distinct pairs of YouTube video categories. Based on Table 1, denote the set of video category pairs that satisfy utility maximization

for the general cost (14) as:

$$C_{GC} = \{(i,j) | i \neq j, i,j \in \{1,5,7,8,9,10,11,12,14,17\}\}$$

We now introduce the three robustness metrics, two for C_{GC} and one for $C \backslash C_{GC}$ as follows:
1. Robustness Metric ($\mathcal{R}_1$) for $C \backslash C_{GC}$: For pairs of video categories that do not satisfy utility maximization, the robustness measure $\mathcal{R}_1$ computes how close these categories are to satisfying utility maximization. Define $\mathcal{R}_1$ for $C \backslash C_{GC}$ as

$$\begin{aligned}
&\mathcal{R}_1(i,j) = \min \frac{|\epsilon|}{\sqrt{(\sum_{x \in X} \sum_{a \in A} u_i(x,a)^2 + u_j(x,a)^2)/2}} \; \forall (i,j) \in C \backslash C_{GC} \text{ s.t.} \\
&\sum_{x \in \mathcal{X}} p_k(x|a)[u_k(x,b) - u_k(x,a)] \leq \epsilon \quad \forall a,b \in \mathcal{A}, \forall k \in \{i,j\} \\
&(G_{i,j} + G_{j,i}) - (G_{i,i} + G_{j,j}) \leq \epsilon,
\end{aligned} \tag{21}$$

where $G_{k,w}$ is as defined in (11). Similar to the perturbation metric used in Varian (1985), note that $\mathcal{R}_1$ is the minimum relaxation for a pair of categories (i,j) to satisfy NIAC and NIAS in (13). Note that $\mathcal{R}_1$ is scale invariant; it is normalized by the average Euclidean norm of the utility vectors for different decision problems.

2. Robustness metric ($\mathcal{R}_2$) for C_{GC}: For pairs of video categories that satisfy utility maximization, what is the minimum perturbation so that they fail the utility maximization test? Define the robustness metric $\mathcal{R}_2$ for C_{GC} as:

$$\begin{aligned}
&\mathcal{R}_2(i,j) = \max \frac{|\epsilon|}{\sqrt{(\sum_{x \in X} \sum_{a \in A} u_i(x,a)^2 + u_j(x,a)^2)/2}} \quad \forall (i,j) \in C_{GC} \text{ s.t.} \\
&\sum_{x \in \mathcal{X}} p_k(x|a)[u_k(x,b) - u_k(x,a)] + \epsilon \leq 0 \quad \forall a,b \in \mathcal{A}, \forall k \in \{i,j\} \\
&(G_{i,j} + G_{j,i}) - (G_{i,i} + G_{j,j}) + \epsilon \leq 0,
\end{aligned}$$

where $G_{k,w}$ is as defined in (11). Note that in contrast to $\mathcal{R}_1$ defined above, $\mathcal{R}_2$ denotes the maximum perturbation/margin for a pair of categories (i,j) such that NIAC and NIAS in (13) still hold. Also $\mathcal{R}_2$ is normalized by the average Euclidean norm of the utility vectors for different decision problems. Alternatively, $\mathcal{R}_2$ can be understood as the minimum perturbation needed for the YouTube category pairs $(i,j) \in C_{GC}$ to fail the utility maximization test for the general cost in Corollary 3.

3. Robustness Metric ($\mathcal{R}_3$) for C_{GC}: Finally, amongst the video category pairs that satisfy general cost utility maximization, how close are they to satisfying the more structured Shannon/Rényi utility maximization? Define the robustness metric $\mathcal{R}_3$ for C_{GC} as:

$$\mathcal{R}_3(i,j) = \frac{1}{|K|} \min \sum_{k=1}^{K} \frac{\|\epsilon_k\|^2}{\|\nabla I^{\beta}(\mu, \alpha_k\|^2} \quad \forall (i,j) \in C_{GC} \text{ s.t.} \tag{22}$$

$$[u_k(x,a), x \in \mathcal{X}, a \in \mathcal{A}]^T = \lambda_{1,k}(\nabla I^{\beta}(\mu, \alpha_k) - \epsilon_k) - \lambda_{2,k}[111..1]^T \quad \forall k \in \{i,j\}, \tag{23}$$

$$\begin{aligned}
&\sum_{x \in \mathcal{X}} p_k(x|a)[u_k(x,b) - u_k(x,a,f)] \leq 0 \quad \forall a,b \in \mathcal{A} \quad \forall k \in \{i,j\}, \\
&(G_{i,j} + G_{j,i}) - (G_{i,i} + G_{j,j}) \leq 0, \\
&\lambda_{1,k} > 0, K = \{i,j\},
\end{aligned}$$

where ϵ_k is a gradient perturbation vector for category k, $G_{k,w}$ is as defined in (11), and (23) is the perturbed version of (19), (20), which in vector form can be written as:

$$[u_k(x,a), x \in \mathcal{X}, a \in \mathcal{A}]^T = \lambda_{1,k}(\nabla I^{\beta}(\mu, \alpha_k)) - \lambda_{2,k}[111..1]^T, \tag{24}$$

where k indexes the k^{th} decision problem. We aim to find the minimum perturbation needed for each category pair in C_{GC} to be utility maximizers with constraints on the information acquisition cost (Shannon and Rényi Mutual Information (18)). For computing $\mathcal{R}_3$, the Euclidean norm of the minimum perturbation vector ϵ_k for each decision problem k is normalized by its gradient vector and averaged over all decision problems.

YouTube Dataset Robustness Analysis Results: With the above three robustness measures, we now analyze the YouTube dataset.

(i) General cost (14): The average $\mathcal{R}_1$ over all category pairs $(i,j) \in C \backslash C_{GC}$ was found to be 1.2×10^{-3}. Moreover, $\mathcal{R}_1 \leq 9 \times 10^{-4}$ for 68% of category pairs $(i,j) \in C \backslash C_{GC}$. These results show that a small perturbation in data for categories in the set $\{2,3,4,6,13,15,16,18\}$ ensures that they satisfy utility maximization for the general cost constraint.

In contrast, the average $\mathcal{R}_2$ over all category pairs $(i,j) \in C_{GC}$ was found to be 7×10^{-3}, approximately 5 times the average value of $\mathcal{R}_1$. $\mathcal{R}_2 \geq 5 \times 10^{-3}$ for 65% of pairs of YouTube categories $(i,j) \in C_{GC}$. This result shows that the minimum perturbation required for category pairs in C_{GC} to fail the utility maximization test is relatively large compared to the minimum perturbation needed for category pairs in $C \backslash C_{GC}$ to pass the test; this highlights that our utility maximization model for the general cost constraint is robust with respect to modeling commenting behavior in YouTube categories.

(ii) Shannon mutual information cost (18): $\mathcal{R}_3$ for $\beta = 1$ over all category pairs $(i,j) \in C_{GC}$ are displayed in Fig. 5. The highest value is 0.096, with 90% of the pairs $(i,j) \in C_{GC}$ having $\mathcal{R}_3 \leq 0.057$. Fig. 5 indicates that all categories in the set $\{1,5,7,8,9,10,11,12,14,17\}$ (56% of categories) are approximately consistent with utility maximization behavior for the Shannon mutual information cost.

(iii) Rényi mutual information cost (18): Fig. 6 displays $\mathcal{R}_3$ averaged over C_{GC}, that is, $\left(\frac{\sum_{(i,j)\in C_{GC}} \mathcal{R}_3(i,j)}{|C_{GC}|}\right)$ for $\beta \in (0,1)$. The maximum averaged $\mathcal{R}_3$ recorded over $\beta \in (0,1)$ was found to be 0.069 which indicates that for any $\beta \in (0,1)$, a small perturbation is sufficient to make user groups for categories in the set $\{1,5,7,8,9,10,11,12,14,17\}$ satisfy utility maximization for Rényi mutual information cost. Interestingly, one can see that the average $\mathcal{R}_3$ decreases with β, implying that the Rényi mutual information cost constraint fits the data better for a higher β, where the average value of $\mathcal{R}_3 \leq 0.01$ for $\beta \geq 0.9$.

According to our robustness analysis, the commenting behavior of user groups in all 18 YouTube video categories approximately satisfies utility maximization for the general cost (14); for 56% of video categories in YouTube, the user groups approximately satisfy rationally inattentive commenting behavior for all information acquisition costs, namely, general cost (14) and Rényi/Shannon mutual information cost (18) for all $\beta \in (0,1]$. The next natural question is: can we predict what type of commenting behavior to expect given the viewcount of a video in a particular category? We explore this aspect in the following sub-section.

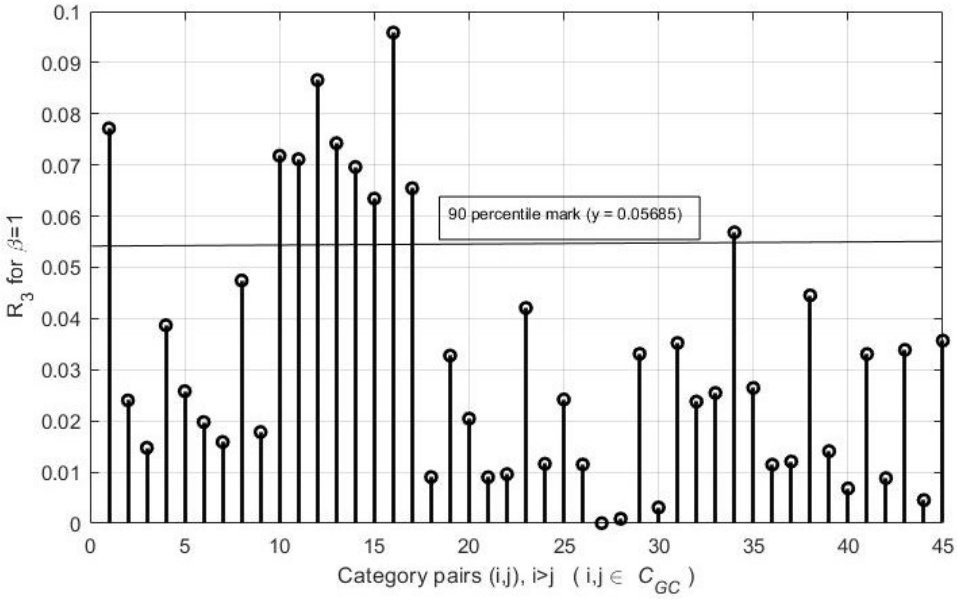

Figure 5: Robustness measure $\mathcal{R}_3(i,j)$ (22) for $\beta = 1$ (Shannon mutual information cost) is plotted on the vertical axis for each category pair $(i,j) \in C_{GC}$. With the 90-th percentile of robustness measure $\mathcal{R}_3$ defined in (22) approximately equal to 0.06, we conclude that all the user groups for categories in the set C_{GC} are approximately consistent with utility maximization for Shannon mutual information cost. This set comprises 90% of videos in the YouTube dataset.

6.3.3 BEHAVIOR PREDICTION USING UTILITY FUNCTION

The results in Sec. 6.2, 6.3.1 and 6.3.2 indicate that YouTube users fit the rationally inattentive Bayesian model remarkably well. The next step is to demonstrate how the rational inattention test in Corollary 3 can be used to predict commenting behavior in specific video categories based on the videos' viewcount (high or low). The main result below is that the comment count (high or low) in the YouTube dataset can be predicted correctly with 83% accuracy.

We divided the YouTube dataset $\mathcal{D}$ into two parts - training data (80%) and testing data (20%). Using Corollary 3, utility estimates for different commenting behavior was constructed for users in each category of YouTube videos in the training data. Out of all categories $(1-18)$, the set $\{1,5,8,9,10,11,12,14\}$ was found to be consistent with utility maximization for the general cost (14); we can construct utility estimates using Corollary 3 for categories belonging to this set. For the remaining categories, it was found from the robustness test in Sec. 6.3.2 that the average minimum relaxation needed for utility maximization behavior is 9.4×10^{-4}.

To quantify the user commenting behavior on the testing data, the *Maximum a posteriori* (MAP) estimate, namely $(\arg\max_{a \in \mathcal{A}} p_k(x|a))$, is computed for each state $x \in \{1,2\}$ in every category k which satisfied the decision test for utility maximization. Note that the MAP estimate is the maximum likelihood estimate of the action unless the prior over actions is uniform. Similarly on the training

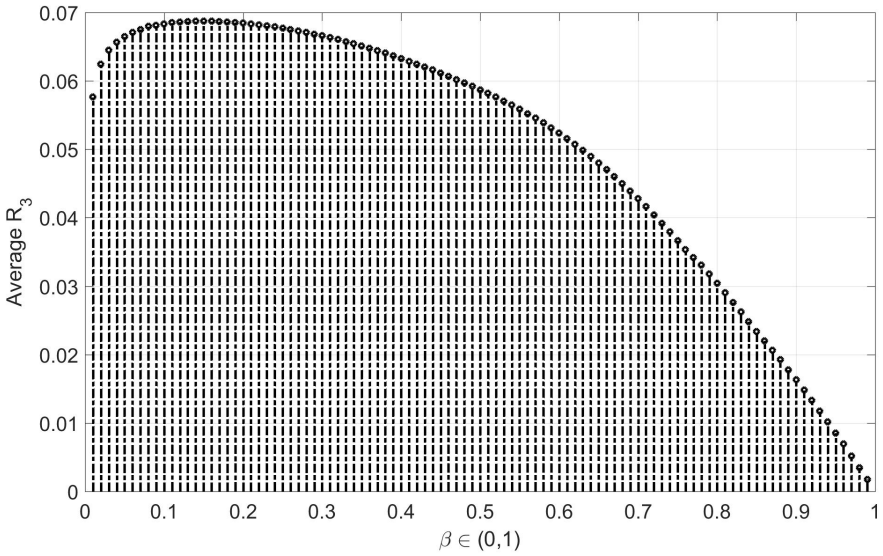

Figure 6: Robustness measure $\mathcal{R}_3$ (22) averaged over C_{GC} for parameter $\beta \in (0,1)$ in Rényi mutual information is plotted on the vertical axis. The average $\mathcal{R}_3$ decreases for $\beta \geq 0.15$ implying that for each category in C_{GC}, a smaller perturbation is needed to satisfy utility maximization model with Rényi mutual information cost for higher β.

data, define the *Maximum Utility Estimate* as $\arg\max_{a \in \mathcal{A}} u_k(x,a)$, for each state $x \in \{1,2\}$ in every category $k \in \{1,5,8,9,10,11,12,14\}$.

Recall from Sec. 6.1 that our actions are numbered from $1-6$, with $1-3$ for low comment count, and the rest for high comment count. We define the error in commenting behavior (action) between prediction (from training data) and observation (from testing data) in each state and category to be the Hamming distance between the MAP estimate (from testing data) and the maximum utility estimate (from the training data) - $|\arg\max_{a \in \mathcal{A}} u_k(x,a) - \arg\max_{a \in \mathcal{A}} p_k(x|a))|$ as shown in Table. 2 [6].

Table 2 also shows that the nature of comment count (high or low) is identical for a particular state $x \in \{1,2\}$ and category k in training and testing data if both MAP estimate and Maximum utility estimate belong to either $\{1,2,3\}$ or $\{4,5,6\}$. From the results in Table 2, atleast one of the two aspects of commenting behavior (nature of comment count (high or low) and sentiment) can be predicted correctly with 83.33% accuracy in 16 (8×2) out of 36 (18×2) sub-categories in YouTube.

6. If the rational inattention cost constraint (3) is omitted from Theorem 2, then *all* 18 *video categories in the YouTube dataset* pass the utility maximization test. However, the predictive capability of this less structured Bayesian utility model is substantially less accurate- the mean squared prediction error (in terms of Hamming distance) is 36% higher (3.85) compared to the original model (2.45) as discussed in Table 2. We thank an anonymous reviewer for suggesting this comparison.

Category	MAP estimate ($x = 1$)	Max. Utility estimate ($x = 1$)	Error ($x = 1$)	MAP estimate ($x = 2$)	Max. Utility estimate ($x = 2$)	Error ($x = 2$)
1	5	5	0	3	1	2
5	5	5	0	5	2	3
8	5	5	0	1	3	2
9	6	5	1	3	2	1
10	5	5	0	3	2	1
11	5	5	0	6	3	3
12	5	4	1	3	2	1
14	6	5	1	3	1	2
17	5	2	3	3	1	2

Table 2: Prediction Results for YouTube Commenting behavior: The mismatch between the commenting behavior in training and testing data is displayed for each state $x \in \{1, 2\}$ (high or low viewcount) and each category which satisfies utility maximization. The mean squared Hamming distance prediction error is 2.45. In 83.33% of these cases, the prediction error is at most 2 units (in terms of Hamming distance) compared to maximum possible error of 5 units. In 27.77% of the cases the prediction error is zero, 61.11% of the cases have an error ≤ 1 unit and 83.33% of the cases have and error ≤ 2 units while the maximum error recorded is 3 units. Thus the rationally inattentive utility maximization model provides an accurate predictor for YouTube commenting behavior.

Discussion

The deep clustering approach in Sec. 6.2 is user centric: it categorizes videos into frames based on the videos' extrinsic framing information. This approach is implemented via Algorithm 2 which preserves "closeness" between videos in both the high dimensional space (where the video thumbnail and description is decoded to) and the low dimensional space (achieved through deep embedded clustering) and achieves a granularity of 4. In comparison, the category dissection approach of Sec. 6.3 is content centric; it exploits differing commenting behavior across YouTube categories (Siersdorfer et al. (2010)) and separates them based on the videos' individual category, achieving a granularity of 18.

Based on extensive analysis of the YouTube dataset, our main conclusions are that users' commenting behavior (comment count and comment sentiment) is i) consistent with rational inattention, ii) depends on the framing information available and the category each video belongs to iii) users prefer to comment on videos that are perceived to be popular. Since the action is also indicative of the perceived sentiment, one can predict how different videos will be perceived based on their popularity over different segments.

Note that our analysis was based on a one-to-one correspondence between the signals and actions which makes the action selection policy exactly the same as the attention function. Relaxing the one-to-one dependency would ensure a higher fraction of data to be consistent with rational inattention. In spite of having used a parsimonious representation of the attention function from the available data, we have unearthed promising insights into YouTube's video interaction dynamics.

That deep embedded clustering adequately captures framing information, that segregating videos by their categories enables the data to fit a rationally inattentive utility maximization model from

which the preference based utility obtained with attention costs rationalizes the YouTube dataset is remarkable. We were also able to predict accurately (83%) the observed commenting behavior from the generated utility function.

7. Conclusion

This paper studied a novel class of inverse reinforcement learning methods for contextual bandit problems with behavioral economics constraints to learn and predict the commenting behavior of YouTube users. The main ideas in this paper involve Bayesian Revealed Preferences, Rational Inattention and Deep Embedded Clustering. Bayesian Revealed Preferences addresses the deeper issue of the existence of a utility function that rationalizes the given data compared to classical inverse reinforcement learning where the existence of such a function is implicitly assumed. Understanding commenting behavior (overall user engagement) in YouTube is important for modeling how humans interactively perceive social multimedia information, for YouTube partner companies to increase their revenue, and for efficient content caching in wireless technologies like 5G.

The main result of this paper was Theorem 2 and Corollary 3 which outline a decision test for utility maximization in Bayesian agents, and Theorem 4 which provides conditions for satisfying additional behavioral economics based information acquisition constraints for the agent. The key application of this paper was to identify rationally inattentive YouTube user groups - groups that were consistent with utility maximization under specific information acquisition cost constraints - and show how the reconstructed utility with rational inattention constraints can be used to predict future commenting behavior in such groups accurately. Our main finding was that YouTube user groups are approximately rationally inattentive; users groups prefer to comment on videos that are popular (high viewcount); and the utility function constructed from the utility maximization decision test can be used to predict commenting behavior of these user groups.

Reproducibility: The computer programs and YouTube datasets needed to completely reproduce all the results in this paper can be obtained from the public GitHub repository: `https://github.com/KunalP117/YouTube_project`.

Acknowledgement

This research was funded in part the U. S. Army Research Office under grant W911NF-19-1-0365, National Science Foundation under grant 1714180, and Air Force Office of Scientific Research under grant FA9550-18-1-0007.

Appendix A. Proofs

A.1 Theorem 2

Proof of necessity of NIAS and NIAC:

1. NIAS (10): Consider any posterior belief $p(x|s)$ such that action $a \in \mathcal{A}$ is chosen by the agent i.e. $\sum_{x\in\mathcal{X}} p(x|s)(u(x,a)-u(x,b)) \geq 0$. The revealed posterior $p(x|a)$ (defined in (10)) is a stochastically garbled version of the actual posterior $p(x|s)$ i.e.,

$$p(x|a) = \sum_{s\in\mathcal{S}(\alpha)} \frac{p(x,a,s)}{p(a)} = \sum_{s\in\mathcal{S}(\alpha)} \frac{p(a|s)p(s)p(x|s)}{p(a)} = \sum_{s\in\mathcal{S}(\alpha)} \frac{p(a,s)}{p(a)} p(x|s) = \sum_{s\in\mathcal{S}(\alpha)} p(s|a)p(x|s)$$

where α is the attention function of the agent defined in (2). Since the optimal action is a, the following holds from optimality for all $a \in \mathcal{A}$

$$\begin{aligned}
&\sum_{x\in\mathcal{X}} p(x|s)(u(x,a)-u(x,b)) \geq 0 \\
\implies &\sum_{s\in\mathcal{S}} p(s|a) \sum_{x\in\mathcal{X}} p(x|s)(u(x,a)-u(x,b)) \geq 0 \\
\implies &\sum_{x\in\mathcal{X}} \left(\sum_{s\in\mathcal{S}} p(s|a)p(x|s)\right)(u(x,a)-u(x,b)) \geq 0 \\
\implies &\boxed{\sum_{x\in\mathcal{X}} p(x|a)(u(x,a)-u(x,b)) \geq 0}
\end{aligned}$$

2. NIAC (11): Let $J(\alpha,\mathcal{A})$ denote the expected utility of the agent when using attention function α (defined in (2)) for decision problem $\mathcal{A}$ (defined in Sec. 6) i.e.,

$$J(\alpha,\mathcal{A}) = \sum_{s\in\mathcal{S}(\alpha)} \left(\sum_{x\in\mathcal{X}} \mu(x)\alpha(s|x)\right) \left[\max_{a\in\mathcal{A}} \sum_{x\in\mathcal{X}} p(x|s)u(x,a)\right]$$

For any sequence of decision problems $\{\mathcal{A}_i\}_{i=1}^k$ and their optimal attention functions $\{\alpha_i\}_{i=1}^k$, we obtain the following relation from optimality using (3)

$$\begin{aligned}
&J(\alpha_j,\mathcal{A}_j) - C(\mu,\alpha_j) \geq J(\alpha_{j+1},\mathcal{A}_j) - C(\mu,\alpha_{j+1}) \\
\implies &J(\alpha_j,\mathcal{A}_j) - J(\alpha_{j+1},\mathcal{A}_j) \geq C(\mu,\alpha_j) - C(\mu,\alpha_{j+1}) \\
\implies &\boxed{\sum_{j=1}^{k} J(\alpha_j,\mathcal{A}_j) - J(\alpha_{j+1},\mathcal{A}_j) \geq \sum_{j=1}^{k} C(\mu,\alpha_j) - C(\mu,\alpha_{j+1}) = 0} \quad (\mathcal{A}_{k+1}=\mathcal{A}_1, \alpha_{k+1}=\alpha_1)
\end{aligned} \tag{25}$$

The following equations establish the relationship between $J(\alpha_i, \mathcal{A}_j)$ and the observed expected utility $G_{i,j}$ defined in (11)

$$\begin{aligned}
G_{i,j} &= \sum_{a\in\mathcal{A}_j}\sum_{x\in\mathcal{X}} p_i(x,a)u_j(x,a) \\
&= \sum_{a\in\mathcal{A}_j}\sum_{s\in\mathcal{S}(\alpha_i)}\sum_{x\in\mathcal{X}} p_i(x,a,s)u_j(x,a) \\
&= \sum_{a\in\mathcal{A}_j}\sum_{s\in\mathcal{S}(\alpha_i)} p_i(a|s)p_i(s)\sum_{x\in\mathcal{X}} p_i(x|s)u_j(x,a) \\
&= \sum_{s\in\mathcal{S}(\alpha_i)} p_i(s)\left[\sum_{a\in\mathcal{A}_j} p_i(a|s)\sum_{x\in\mathcal{X}} p_i(x|s)u_j(x,a)\right] \\
&\leq \sum_{s\in\mathcal{S}(\alpha_i)} p_i(s)\left[\max_{a\in\mathcal{A}_j}\sum_{x\in\mathcal{X}} p_i(x|s)u_j(x,a)\right] = J(\alpha_i,\mathcal{A}_j) \\
\implies & J(\alpha_i,\mathcal{A}_j) \geq G_{i,j}
\end{aligned} \tag{26}$$

Clearly, equality holds above if $i = j$. Using (26), we further write (25) as

$$\sum_{j=1}^{k} J(\alpha_j,\mathcal{A}_j) \geq J(\alpha_{j+1},\mathcal{A}_j) \implies \sum_{j=1}^{k} G_{j,j} \geq \sum_{i=1}^{k} J(\alpha_{j+1},\mathcal{A}_j) \geq \sum_{i=1}^{k} G_{j,j+1}$$

Hence, $\boxed{\sum_{j=1}^{k} G_{j,j} - G_{j,j+1} \geq 0}$ $\quad (G_{k,k+1} = G_{k,1})$

Proof of sufficiency of NIAS and NIAC:
Given any decision problem, our assumption of a one-to-one mapping from the set of signals ($\mathcal{S}(\alpha)$) to the set of actions ($\mathcal{A}$) in Sec. 5 implies $p(s|a) = 1$ for the signal corresponding to the action $a \in \mathcal{A}$ and 0 elsewhere. Thus, we first show that the NIAS condition (10) in Theorem 1 implies (2) in Definition 1.

$$\begin{aligned}
\text{NIAS: } & \sum_{x\in\mathcal{X}} p(x|a)\left[u(x,a) - u(x,b)\right] \geq 0 \quad \forall a,b \in \mathcal{A} \\
&= \sum_{s\in\mathcal{S}(\alpha)}\sum_{x\in\mathcal{X}} \frac{p(a,s,x)}{p(a)}\left[u(x,a) - u(x,b)\right] \geq 0 \quad \forall a,b \in \mathcal{A} \\
&= \sum_{s\in\mathcal{S}(\alpha)} \frac{p(a|s)p(s)}{p(a)}\left[\sum_{x\in\mathcal{X}} p(x|s)\left[u(x,a) - u(x,b)\right]\right] \geq 0 \quad \forall a,b \in \mathcal{A} \\
&= \sum_{s\in\mathcal{S}(\alpha)} p(s|a)\left[\sum_{x\in\mathcal{X}} p(x|s)\left[u(x,a) - u(x,b)\right]\right] \geq 0 \quad \forall a,b \in \mathcal{A} \\
&= \sum_{x\in\mathcal{X}} p(x|s_a)\left[u(x,a) - u(x,b)\right] \geq 0 \quad \forall a,b \in \mathcal{A} \quad (p(s_a|a) = 1)
\end{aligned}$$

Define a square matrix $A_{K\times K}$ with elements $A_{i,j} = G_{i,j}$ (defined in (11)). The NIAC condition (11) implies that the cumulative expected utility across all K decision problems is maximized by employing action selection policy $\pi_k(a|x)$ to the k^{th} decision problem. This is identical to the allocation problem in Koopmans and Beckmann (1957) with A defined above as the profitability matrix and the optimal allocation function mapping plants to locations being the identity map. Koopmans and Beckmann (1957) solve the allocation

problem by defining an equivalent linear program. From duality theory (Boyd and Vandenberghe (2004)), there exist shadow prices $C_k \in \mathbb{R}^+$ for each attention selection policy $\pi_k(a|x)$ such that

$$\begin{aligned} & G_{j,i} - G_{j,j} \leq C_i - C_j \quad \forall i, j \in \{1, 2 \dots K\} \\ \implies & G_{i,i} - C_i \leq G_{i,j} - C_j \end{aligned}$$

Note that under the parsimonious representation assumption in Sec. 5, the attention function α_i can be replaced with the action selection policy π_i defined in (4). Using the shadow prices C_i, we construct the cost function C defined in (2) as follows

$$C(\mu, \alpha) = \begin{cases} C_i, & \exists i \in \{1, 2 \dots K\} \text{ s.t } \alpha = \alpha_i \\ \infty & \text{otherwise} \end{cases} \tag{27}$$

Clearly, the above construction of the information cost function $C(\mu, \alpha)$ implies that the attention selection policies are optimally chosen over decision problems, hence proving (3) in Definition 1.

A.2 Theorem 4

In the context of the convex optimization problem defined in (17) (Boyd and Vandenberghe (2004), pg. 121), the following KKT conditions are necessary and sufficient for the global optimum $p^*(x, a)$ to satisfy, provided strong duality (for eg. Slater's condition) holds:

$$\begin{aligned} & \exists \lambda_1^*, \lambda_2^* \in \mathbb{R} \; s.t \\ & \nabla \left(\sum_{\mathcal{X}, \mathcal{A}} p^*(x, a) u(x, a) \right) - \lambda_1^* \nabla (I_\beta(\mu, \alpha^*) - \kappa_{max}) + \lambda_2^* \nabla (\mu - [\{p^*(x), x \in \mathcal{X}\}]) = 0 \\ & p^*(x) = \mu(x), \; \forall x \in \mathcal{X} \quad I_\beta(\mu, \alpha^*) \leq \kappa_{max} \\ & \lambda_1^* \geq 0, \quad \lambda_1^* \left(I_\beta(\mu, \alpha^*) - \kappa_{max} \right) = 0. \end{aligned}$$

Slater's condition can be checked by setting $p(x) = \mu(x), p(a|x) = p_{arb}(a)$, where $p_{arb}(a)$ is any arbitrary probability mass function over the set of actions $\mathcal{A}$. This particular choice of the joint probability mass function $p(x, a)$ results in $I_\beta(\mu, \alpha) = 0 < \kappa_{max}, \forall \beta \in (0, 1]$ thus satisfying the Slater's condition.

Let $p^*_{nc}(x, a) \in \arg\max_{p(x,a)} \left(\sum_{\mathcal{X}, \mathcal{A}} p^*(x, a) u(x, a) \right)$ denote the unconstrained optimum for the convex optimization problem (17) and denote its corresponding attention function (defined in (2)) to be α^*_{un}. Then, setting $\kappa_{max} < I_\beta(\mu, \alpha^*_{un})$ implies $\lambda_1^* > 0$, where λ_1^{ast} is the Lagrange multiplier corresponding to the inequality constraint $I_\beta(\mu, \alpha) \leq \kappa_{max}$. Define the Lagrangian $L(p) = \left(\sum_{\mathcal{X}, \mathcal{A}} p(x, a) u(x, a) \right) - \lambda_1^* (I_\beta(\mu, \alpha) - \kappa_{max}) + \lambda_2^* (\mu - [\{p(x), x \in \mathcal{X}\}])$.

1. Consider the Renyi mutual information defined in (15) with $\beta \in (0,1)$:

$$\frac{\partial}{\partial p(\hat{x},\hat{a})}(L(p)) = u(\hat{x},\hat{a}) - \lambda_1^* \frac{\partial}{\partial p(\hat{x},\hat{a})}(I_\beta(\mu,\alpha)) + \lambda_2^* = 0$$

$$\frac{\partial}{\partial p(\hat{x},\hat{a})}(I_\beta(\mu,\alpha)) = \frac{\partial}{\partial p(\hat{x},\hat{a})}\left[\frac{1}{\beta-1}\ln\left(\sum_{x\in\mathcal{X}}\sum_{a\in\mathcal{A}}\frac{p^\beta(x,a)}{\mu^{\beta-1}(x)p^{\beta-1}(a)}\right)\right]$$

$$= \frac{1}{(\beta-1)\left(\mathbb{E}\left[\frac{p(x|a)}{\mu(x)}^{\beta-1}\right]\right)}\left[-(\beta-1)p^{-\beta}(\hat{a})\sum_{x\in\mathcal{X}}\frac{p^\beta(x,\hat{a})}{\mu^{\beta-1}(x)} + p^{1-\beta}(\hat{a})\frac{\beta p^{\beta-1}(\hat{x},\hat{a})}{\mu^{\beta-1}(\hat{x})}\right]$$

$$= \frac{1}{(\beta-1)\left(\mathbb{E}\left[\frac{p(x|a)}{\mu(x)}^{\beta-1}\right]\right)}\left[-(\beta-1)p^{1-\beta}(\hat{a})\mathbb{E}_x\left[\eta^{\beta-1}(x,\hat{a})\right] + \beta p^{1-\beta}(\hat{a})\eta^{\beta-1}(\hat{x},\hat{a})\right]$$

$$= \frac{\beta\eta^{\beta-1}(\hat{x},\hat{a}) - (\beta-1)\mathbb{E}_x\left[\eta^{\beta-1}(x,\hat{a})\right]}{(\beta-1)(\mathbb{E}\left[\eta^{\beta-1}(x,a)\right])p^{\beta-1}(\hat{a})}$$

$$\therefore \boxed{\frac{\partial}{\partial p(\hat{x},\hat{a})}(L(p)) = u(\hat{x},\hat{a}) - \lambda_1^*\frac{\beta\eta^{\beta-1}(\hat{x},\hat{a}) - (\beta-1)\mathbb{E}_x\left[\eta^{\beta-1}(x,\hat{a})\right]}{(\beta-1)(\mathbb{E}\left[\eta^{\beta-1}(x,a)\right])p^{\beta-1}(\hat{a})} + \lambda_2^*}$$

2. Consider the Shannon mutual information defined in (15) with $\beta = 1$:

$$\frac{\partial}{\partial p(\hat{x},\hat{a})}(L(p)) = u(\hat{x},\hat{a}) + \lambda_1^* \frac{\partial}{\partial p(\hat{x},\hat{a})}(I(\mu,\alpha)) - \lambda_2^* = 0$$

$$\frac{\partial}{\partial p(\hat{x},\hat{a})}(I(\mu,\alpha)) = \frac{\partial}{\partial p(\hat{x},\hat{a})}\left(\underbrace{H(\mu)}_{\text{constant}} + \sum_{x\in\mathcal{X}}\sum_{a\in\mathcal{A}} p(x,a)\ln(p(x,a)/p(a))\right)$$

$$= \frac{\partial}{\partial p(\hat{x},\hat{a})}\left(\sum_{x\in\mathcal{X},a\in\mathcal{A}} p(x,a)\ln(p(x,a)) - \sum_{a\in\mathcal{A}} p(a)\ln(p(a))\right)$$

$$= p(\hat{x},\hat{a}).\frac{1}{p(\hat{x},\hat{a})} + \ln(p(\hat{x},\hat{a})) - p(\hat{a}).\frac{1}{p(\hat{a})} - \ln p(\hat{a}) = \ln p(\hat{x}|\hat{a})$$

$$\therefore \boxed{\frac{\partial}{\partial p(\hat{x},\hat{a})}(L(p)) = u(\hat{x},\hat{a}) - \lambda_1^*\ln p(\hat{x}|\hat{a}) + \lambda_2^*}$$

Appendix B. Finite Sample Performance Analysis of the Agent's Action-Selection Policy

This appendix gives a finite sample analysis of the agent's action selection policy defined in Sec. 2. Since the results are somewhat tangential to our main application, we have put this analysis in an appendix. In YouTube, given the action selection policy (4), the observer (data analyst) can construct an optimal recommender system for agents' commenting behavior that is consistent with the agent's commenting preferences while ensuring the agent's commenting behavior satisfies rational inattention. The construction of the optimal policies $\boldsymbol{\pi}(a|x,f)$ is based on a variance-penalized optimization method using finite sample bounds on the total expected utility (12).

Consider the maximum likelihood estimate of the agent's action-selection policy $\hat{\pi}(a|x,f)$ (6). An important question related to performance analysis of these estimators is: How far is the net utility obtained using this estimated policy (based on a finite dataset) compared to the actual net utility $V(\boldsymbol{\pi}(a|x,f))$ (12) which uses the true policy $\boldsymbol{\pi}(a|x,f)$? Using an extension of the empirical Bernstein inequality to the space of

continuous function classes

$$\mathcal{F}_{\Pi} = \{f_{\pi,k} : \mathcal{X} \times \mathcal{A}_k \times N \to [0,1]\},\ f_{\pi,k} = M\frac{\pi_k(a|x,f)}{\hat{\pi}_k(a|x,f)}u(x,a,f) = M\bar{u}(\pi_k(a|x,f)) \tag{28}$$

we can construct a finite sample bound between the observed net utility $V(\hat{\boldsymbol{\pi}}(a|x,f))$ and an estimate of the net utility $V(\boldsymbol{\pi}(a|x,f))$ for the unobserved policy $\boldsymbol{\pi}(a|x,f)$. In (28), M is a normalization constant which ensures $f_{\pi,k} \in [0,1]$, $\hat{\pi}_k(a|x,f)$ is the observed policy (6), and $\pi_k(a|x,f)$ is an unobserved policy. By bounding the function class (28) using the uniform covering number and employing the double-sampling method (Anthony and Bartlett (2009)), Theorem 5 results.

Theorem 5 *Let $\bar{u}(\pi_k)$ be a random variable with T_k independent and identically distributed (i.i.d.) samples in $\mathcal{D}$. Then with probability $1-\gamma$ the random vector $(a_t, x_t) \sim \pi_k$, for a stochastic hypothesis class $\pi_k \in \Pi$, $T_k \geq 16$, and $\lambda = \sqrt{18\ln(10\mathcal{N}_{\infty}\{1/T_k, \mathcal{F}_{\Pi}, 2T_k\}/\gamma)}$, satisfies*

$$V(\pi_k) \leq \hat{V}(\pi_k) + \lambda\sqrt{\frac{\mathrm{Var}[\bar{u}(\pi_k)]}{T_k}} + \frac{15\lambda^2}{18M(T_k-1)} \tag{29}$$

where $\mathcal{N}_{\infty}\{1/T_k, \mathcal{F}_{\Pi}, 2T_k\}/\gamma)$ is the uniform covering number.

Theorem 5 provides a probabilistic bound between the estimated net utility $\hat{V}(\pi_k)$ and actual net utility $V(\pi_k)$ that only depends on the dataset $\mathcal{D}$ and the coefficient λ. Therefore, for constructing the true policy $\boldsymbol{\pi}(a|x,f)$, one would maximize the net utility $\hat{V}(\pi_k)$ while minimizing the variance term with a coefficient $\bar{\lambda} \geq 0$. Note that in Theorem 5, λ encodes the entropy of the function class $\mathcal{F}_{\Pi}$, which is dependent on the number of samples T_k, uniform covering number $\mathcal{N}_{\infty}\{\cdot\}$, and γ which is a measure of the confidence of the estimate. For the function class (28), $\mathcal{N}_{\infty}\{\cdot\}$ is polynomial in the sample size T_k (Maurer and Pontil (2009); Vapnik and Chervonenkis (2015); Sauer (1972))–this ensures as the sample size increases that $\widehat{V}(\pi_k) \to V(\pi_k)$. Using Theorem 5, the convex feasibility problem

$$\begin{aligned}
&\boldsymbol{\pi}(a|x,f) \in \underset{\pi_k \in \Pi}{\arg\max} \left\{ \sum_{k=1}^{K} V(\boldsymbol{\pi}_k(a|x,f)) - \bar{\lambda}_k\sqrt{\frac{\mathrm{Var}[u(a,x,f)]}{T_k}} \right\} \\
&\text{s.t.} \quad \sum_{a \in \mathcal{A}_k} \pi_k(a|x,f) = 1, \quad \pi_k(a|x,f) \geq 0 \\
&\mathcal{L}(u(a,x,f), \hat{\pi}_k(a|x,f)) \leq 0 \\
&\mathcal{L}(u(a,x,f), \pi_k(a|x,f)) \leq 0 \quad \forall x \in \mathcal{X}, \forall a \in \mathcal{A}_k, \forall k \in \{1,\ldots,K\}, \forall f \in \{1,\ldots,N\} \qquad (30) \\
&\text{where } \mathcal{L}(\cdot) \text{ refers to the inequalities (13) in Corollary 3.} \qquad (31)
\end{aligned}$$

can be used to construct the optimal policy $\pi_k(a|x,f)$ that maximizes the net utility $V(\boldsymbol{\pi}(a|x,f))$ while ensuring the policy is consistent with rational inattention. The regularization term $\bar{\lambda}_k$ in (31) balances the maximization of the net utility $V(\boldsymbol{\pi}(a|x,f))$ while accounting for the finite-sample variance associated with estimating $V(\boldsymbol{\pi}(a|x,f))$ for policies $\boldsymbol{\pi}(a|x,f)$ that are different from $\hat{\boldsymbol{\pi}}(a|x,f)$. The lower the value of $\bar{\lambda}_k$, the more risk-seeking the generated optimal policy.

As seen, the objective in (31) is based on the finite-sample bound provided in Theorem 5. An important property of (31) is that it provides a method to construct optimal policy recommendations for agents. Specifically, in a state x and frame f recommendations can be tuned such that the probability of selecting action a is consistent with the optimal policy $\boldsymbol{\pi}(a|x,f)$ to maximize the agent's total expected utility V in (12) without impacting the preferences of the agent.

Appendix C. Denoising Autoencoder Architecture for YouTube Title and Thumbnail

Algorithm 1 (in the main text) outlined the steps in the deep embedding method for constructing the preference invariant frames. Algorithm 2 below describes the details of the deep clustering procedure explained in Sec. 3.

The denoising autoencoder is comprised of stacked long short term memory (LSTM) and convolutional neural network (CNN) which are detailed in Sec. C.1 and Sec. C.2. To ensure the denoising autoencoder is robust to variations in the title and thumbnail input (e.g. good generalization performance), we introduce noise into the input training data. Possible methods to introduce noise into the network include using drop-out (Srivastava et al. (2014)) and drop-path (Huang et al. (2016)) methods. Here we apply Gaussian noise to the input images and numeric representation of the words, and additionally include drop-out layers in the LSTM and CNN networks.

Algorithm 2 Deep Embedded Clustering for Framing Association

Require: Set of framing information $\{f_t\}_{t=1}^T$, number of unique frames N, stopping threshold $\delta \in (0,1)$, confidence threshold $\delta_c \in (0,1)$, and updating interval ζ.

PRE-TRAIN

Pre-train the denoising autoencoder without any frame association.

return n_t^0 (defined in Sec. 6.2)

INITIALIZE

Initialize the N cluster centers Ψ^o using k-means clustering in the latent space and set $\varepsilon = 0$.

DEEP CLUSTERING

Train the deep clustering autoencoder and frame association layers to iteratively generate n_t^i.

$i = 0$

while $\sum_t \mathbb{1}\{n_t^0 \neq n_t^i\} \geq T\delta$ **do**

if $i \mod \zeta == 0$ **then**

Compute all latent points $\{z_t = r(w(f_t))\}_{t=1}^T$

Compute P using (9)

Set $n_t^o \leftarrow n_t^i$ for all $t = 1, 2, \ldots T$

Compute new cluster labels $n_t^i = \arg\max_{n \in \{1,\ldots,N\}} \{q_{t,n}\}$.

else

Select mini-batch sample from $\{f_t\}_{t=1}^T$ and update the weights of the autoencoder and frame association layers to minimize the loss (7).

end if

$i = i + 1$

end while

return Invariant frames $n_t^i \quad \forall t \in \{1, \ldots, T\}$ such that $\max_n \{q_{t,n}\} > \delta_c$.

C.1 Text Processing of the YouTube Title

The design of autoencoders for text data is challenging as a result of the power-law distribution of words (Mandelbrot (1953)) and the long-range dependencies (grammars) between words. To address these challenges, we use previously constructed word embeddings to convert the words into a numeric vector. We then employ a LSTM networks for the encoder and decoder blocks of the autoencoder which focus on text processing. The combination of using word embeddings and LSTMs allows the network to use prior knowledge of similar words while simultaneously learning how to cluster similar sentences into a unique frame.

Prior to transforming the words into their numeric embedding, we apply a lemmatization transformation. Lemmatization reduces the number of variations of words necessary to consider as it groups all the inflected forms a word into a single base representation. For example, the verb "to walk" may appear as "walk", "walked", "walks", "walking" which are all converted to "walk" via the lemmatization transformation. To

perform the lemmatization transformation we use the WordNet lemmatizer [7]. The WordNet lemmatizer is comprised of two resources, a set of rules which identify the inflectional endings that can be detached from individual words, and a list of exceptions for irregular word forms. WordNet first checks the exceptions, then remove any inflectional endings from the words. Having performed the lemmatization operation, we now construct numeric vector representations of the words. A popular method to perform this task is to use distributed representations of words (e.g. word embeddings). The distributed representation of words in a vector space are designed such that words with similar semantic meaning have similar latent space representations. Equivalently, words with similar meaning will cluster together in the word embedding space. Two popular word embeddings are the Word2Vec (Mikolov et al. (2013)) and Glove (Pennington et al. (2014)) models. For the clustering algorithm we use the Glove embedding that was constructed using over 2 billion tweets and is comprised of over 1.2 million words. The possible dimension of the word embedding space is 25, 50, 100, or 200. Here we use a word embedding dimension of 25.

Given the word embeddings of the sentence $w(f)$, we use an LSTM encoder-decoder framework to learn latent space representations of the titles (Goldberg (2016); Sutskever et al. (2014); Goodfellow et al. (2016); Géron (2017)). To construct the latent space representation of the sentences, we use a stacked LSTM architecture. Note that stacked LSTMs are able to capture grammatical information in the title at different scales. It was illustrated in Goldberg (2016); Sutskever et al. (2014) that stacked LSTMs tend to have superior predictive performance compared to single layer LSTMs for natural language processing tasks.

C.2 Image Processing of the YouTube Thumbnail

In the denoising autoencoder, image processing is performed using a VGG (Visual Geometry Group) based architecture (Vedaldi and Lenc (2015)). Given the latent space representation z_t from the encoder, the image decoder is used to reconstruct the original input image. To perform this task requires the use of deconvolution and upsampling layers. However, deconvolution layers are not used in CNN autoencoders. Instead a mixture of convolutional and upsampling layers are employed. In the most extreme case, a single upsampling layer can be used to directly reconstruct the images from the latent space as illustrated in Long et al. (2015). A commonly used method is to construct multiple transposed convolution (also known as fractionally strided convolutions) layers in combination with upsampling layers. Using the transposed convolution layers instead of the standard convolution layers ensures that "checkerboard" artifacts are removed from the decoded image (Odena et al. (2016)).

Appendix D. User Group Statistics in YouTube Dataset

In this appendix, we provide additional information about the YouTube dataset analyzed in this paper. Figure 7 lists each video category along with the total number of views. Note that the video categories "Unavailable" or "Removed" are videos flagged by YouTube as being suspected of violating YouTube's video policies[8].

Based on YouTube parameter description in Sec. 6.1, each frame (indexed as $1-4$) in Sec. 6.2 comprises two decision problems- the first problem for videos in gaming category and the second problem for videos belonging to non-gaming categories. We subdivide the videos in our YouTube dataset into 8 sub-categories: $1-4$ and $5-8$ correspond to videos belonging to gaming and non-gaming categories in frames $1-4$ respectively. The average number of interacting users for videos in sub-categories $1-4$ ranges from 493 to 1133 users; the average number of interacting users for videos in sub-categories $5-8$ ranges from 368 to 513 users.

Similarly in Sec. 6.3, each video category (indexed from $1-18$) is associated with a distinct YouTube user group. We again subdivide the videos in our YouTube dataset into 36 sub-categories where sub-categories $1-18$ and $19-36$ correspond to videos in each video category ($1-18$) with a high viewcount (greater than 10000 views) and low viewcount (lesser than 10000 views) respectively. The average number of interacting

7. https://wordnet.princeton.edu/wordnet/

8. Refer to `https://www.youtube.com/yt/about/policies/#community-guidelines` for details

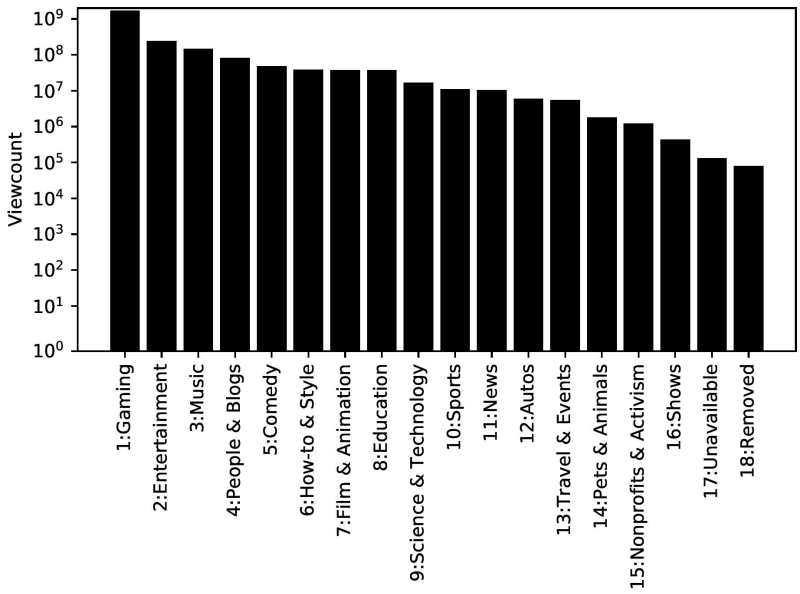

Figure 7: Viewcount summed over all videos (vertical axis) of the 18 video categories of the YouTube dataset $\mathcal{D}$. The 18 categories are listed on the horizontal axis.

users for videos in sub-categories $1-18$ ranges from 149 to 4596 users; the average number of interacting users for videos in sub-categories $19-36$ ranges from 8 to 1801 users.

Appendix E. Estimating the Agent's Attention Function and Choice Function

If the dataset $\mathcal{D}$ satisfies rational inattention, the observer (analyst) can estimate the agent's attention function $\alpha_k(s|x)$ and choice function $\eta_k(a|s)$.

Constructing the agent's attention function $\alpha_k(s|x)$ and choice function $\eta_k(a|s)$ which were defined in (2) and (4) requires the posterior distribution $p_k(x|a)$. First, consider the signal set $\mathcal{S}(\alpha_k)$ of all observed posterior state distributions of the agent for attention function $\alpha_k(s|x)$ using

$$\mathcal{S}(\alpha_k) = \{p_k(x|a) : a \in \mathcal{A}_k\}, \quad p_k(x|a) = \frac{\mu(x)\pi_k(a|x)}{\sum_{y\in\mathcal{X}} \mu(y)\pi_k(a|y)}. \tag{32}$$

Each posterior distribution $p_k(x|a)$ is associated with a single signal $s \in \mathcal{S}(\alpha_k)$. The posterior distribution $p_k(x|a)$ in (32) is equal to the true posterior distribution $p_k(x|s)$ in (1) only if the choice function $\eta_k(a|s)$ produces a single action $a \in \mathcal{A}_k$ for each $s \in \mathcal{S}(\alpha_k)$ with probability one. Otherwise the posterior distribution $p_k(x|a)$ is given by the weighted sum

$$p_k(x|a) = \frac{\sum_{s\in\mathcal{S}(\alpha_k)} \eta_k(a|s)p_k(x|s)p_k(s)}{\sum_{x\in\mathcal{X}}\sum_{s\in\mathcal{S}(\alpha_k)} \eta_k(a|s)p_k(x|s)p_k(s)}. \tag{33}$$

Note that without explicit knowledge of the choice and attention functions of the agent, the stochastic choice dataset can not be used to determine if $p_k(x|a) = p_k(x|s)$. Having $p_k(x|a) = p_k(x|s)$ is not required to determine if the agent satisfies rational inattention.

Given $p_k(x|a)$, for each signal $s \in \mathcal{S}(\alpha_k)$, the associated attention function is

$$\alpha_k(s|x) = \sum_{a\in\mathcal{A}_k} \eta_k(a|s)\alpha_k(s|x) = \sum_{a\in\mathcal{A}_k} \pi_k(a|x)\mathbf{1}\{p_k(x|a) = s\} \tag{34}$$

where the second equality results from using the data matching expression in (4). Note that (34) is only equal to the agent's attention function $\rho_k(r|x)$ if the observed and true posterior distributions are equal. If $\rho_k(r|x)$ is the true attention function then

$$\alpha_k(s|x) = \sum_{r \in \mathcal{S}(\rho_k)} \sum_{a \in \mathcal{A}_k} \eta_k(a|r)\rho_k(r|x)\mathbf{1}\{p_k(x|a) = s\}. \tag{35}$$

It must be the case that the observed attention function $\alpha_k(s|x)$ is weakly less informative than the true attention function $\rho_k(r|x)$. Equivalently, the observed attention function is a noisy version of the true attention function. Theorem 1 however does not require we know the true attention function $\rho_k(r|x)$ of the agent to test if the agent's behavior satisfies rational inattention.

The observed choice function of the agent is given by

$$\eta_k(a|s) = \frac{\sum_{x \in \mathcal{X}} \mu(x)\pi_k(a|x)}{\sum_{b \in \mathcal{A}_k} \sum_{x \in \mathcal{X}} \mu(x)\pi_k(b|x)\mathbf{1}\{p_k(x|b) = s\}} \tag{36}$$

which is the ratio of the number of times action $a \in \mathcal{A}_k$ was selected over all other possible actions $b \in \mathcal{A}_k$ for the prior distribution $s \in \mathcal{S}(\alpha_k)$. The observed choice function provides no information on the true choice function over the posterior distributions $r \in \Gamma(\rho_k)$ that result from the true attention function unless the actual and observed posterior distributions are equal. Note however that the observed attention function $\alpha_k(s|x)$ (34) and choice function $\eta_k(a|s)$ (36) are consistent with the agent's observed action-selection policy $\pi_k(a|x)$ as expressed in (4).

References

S. N. Afriat. The construction of utility functions from expenditure data. *International economic review*, 8(1): 67–77, 1967.

S. Alhabash, J. Baek, C. Cunningham, and A. Hagerstrom. To comment or not to comment?: How virality, arousal level, and commenting behavior on youtube videos affect civic behavioral intentions. *Computers in human behavior*, 51:520–531, 2015.

M. Anthony and P. Bartlett. *Neural network learning: Theoretical foundations*. cambridge university press, 2009.

A. Aprem and V. Krishnamurthy. Utility change point detection in online social media: A revealed preference framework. *IEEE Transactions on Signal Processing*, 65(7), April 2017.

A. Badanidiyuru, J. Langford, and A. Slivkins. Resourceful contextual bandits. In *Conference on Learning Theory*, pages 1109–1134, 2014.

Y. Bengio. Learning deep architectures for AI. *Foundations and trends in Machine Learning*, 2(1):1–127, 2009.

S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

A. Caplin and M. Dean. Revealed preference, rational inattention, and costly information acquisition. *The American Economic Review*, 105(7):2183–2203, 2015.

A. Caplin and D. Martin. A testable theory of imperfect perception. *The Economic Journal*, 125(582):184–202, 2015.

W. Diewert. Afriat's theorem and some extensions to choice under uncertainty. *The Economic Journal*, 122 (560):305–331, 2012.

S. Frintrop, E. Rome, and H. I. Christensen. Computational visual attention systems and their cognitive foundations: A survey. *ACM Transactions on Applied Perception (TAP)*, 7(1):1–39, 2010.

J. Fu, K. Luo, and S. Levine. Learning robust rewards with adversarial inverse reinforcement learning. *arXiv preprint arXiv:1710.11248*, 2017.

A. Géron. Hands-on machine learning with scikit-learn and tensorflow: concepts, tools, and techniques to build intelligent systems, 2017.

A. Glöckner and T. Betsch. Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *MPI Collective Goods Preprint*, 2(2008), 2008.

Y. Goldberg. A primer on neural network models for natural language processing. *J. Artif. Intell. Res.(JAIR)*, 57:345–420, 2016.

I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT press, 2016.

X. Guo, L. Gao, X. Liu, and J. Yin. Improved deep embedded clustering with local structure preservation. In *International Joint Conference on Artificial Intelligence (IJCAI-17)*, pages 1753–1759, 2017.

D. Hadfield-Menell, S. Milli, P. Abbeel, S. Russell, and A. Dragan. Inverse reward design. In *Advances in Neural Information Processing Systems*, pages 6768–6777, 2017.

M. M. Hayhoe, A. Shrivastava, R. Mruczek, and J. B. Pelz. Visual memory and motor planning in a natural task. *Journal of vision*, 3(1):6–6, 2003.

C. Helmers, P. Krishnan, and M. Patnam. Attention and saliency on the internet: Evidence from an online recommendation system. *CEPR Discussion Paper No. DP10939*, 2015.

S. Ho and S. Verdú. Convexity/concavity of rényi entropy and α-mutual information. In *Information Theory (ISIT), 2015 IEEE International Symposium on*, pages 745–749. IEEE, 2015.

W. Hoiles, O. Namvar, V. Krishnamurthy, N. Dao, and H. Zhang. Caching in the youtube content distribution network: A revealed preference game-theoretic learning approach. *IEEE Transactions on Cognitive Communications and Networking*, 1(1):71–85, 2015.

W. Hoiles, A. Aprem, and V. Krishnamurthy. Engagement and popularity dynamics of youtube videos and sensitivity to meta-data. *IEEE Transactions on Knowledge and Data Engineering*, 29(7):1426–1437, 2017.

G. Huang, Y. Sun, Z. Liu, D. Sedra, and K. Weinberger. Deep networks with stochastic depth. In *European Conference on Computer Vision*, pages 646–661. Springer, 2016.

L. Huang and H. Liu. Rational inattention and portfolio selection. *The Journal of Finance*, 62(4):1999–2040, 2007.

T. Judd, F. Durand, and A. Torralba. A benchmark of computational models of saliency to predict human fixations, 2012.

R. E. Kalman. When is a linear control system optimal? *Journal of Basic Engineering*, pages 51–60, 1964.

L. Khan. Social media engagement: What motivates user participation and consumption on youtube? *Computers in Human Behavior*, 66:236–247, 2017a.

M. L. Khan. Social media engagement: What motivates user participation and consumption on youtube? *Computers in Human Behavior*, 66:236–247, 2017b.

B. Kim and J. Pineau. Socially adaptive path planning in human environments using inverse reinforcement learning. *International Journal of Social Robotics*, 8(1):51–66, 2016.

T. C. Koopmans and M. Beckmann. Assignment problems and the location of economic activities. *Econometrica: journal of the Econometric Society*, pages 53–76, 1957.

I. Krajbich, C. Armel, and A. Rangel. Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13(10):1292, 2010.

H. Kwon and A. Gruzd. Is offensive commenting contagious online? examining public vs interpersonal swearing in response to donald trump?s youtube campaign videos. *Internet Research*, 27(4):991–1010, 2017.

P. G. Lange. Publicly private and privately public: Social networking on youtube. *Journal of computer-mediated communication*, 13(1):361–380, 2007.

D. Lee, R. J. Kauffman, and M. E. Bergen. Image effects and rational inattention in internet-based selling. *International Journal of Electronic Commerce*, 13(4):127–166, 2009.

H. Levy and H. M. Markowitz. Approximating expected utility by a function of mean and variance. *The American Economic Review*, 69(3):308–317, 1979.

J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.

M. Lopes, F. Melo, and L. Montesano. Active learning for reward estimation in inverse reinforcement learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 31–46. Springer, 2009.

B. Mandelbrot. An informational theory of the statistical structure of language. *Communication theory*, 84: 486–502, 1953.

S. Mannor and J. N. Tsitsiklis. Mean-variance optimization in markov decision processes. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, pages 177–184, 2011.

J. G. March. Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, pages 587–608, 1978.

D. Martin. Strategic pricing with rational inattention to quality. *Games and Economic Behavior*, 104:131–145, 2017.

F. Matejka and A. McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98, 2015.

A. Maurer and M. Pontil. Empirical bernstein bounds and sample variance penalization. In *The 22nd Conference on Learning Theory*, 2009.

J. L. McClelland and D. E. Rumelhart. *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. MIT press, 1989.

T. Mikolov, O. Plchot, O. Glembek, L. Burget, and J. Cernockỳ. Pca-based feature extraction for phonotactic language recognition. In *Odyssey*, page 42, 2010.

T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.

A. Y. Ng, S. J. Russell, et al. Algorithms for inverse reinforcement learning. In *Icml*, volume 1, page 2, 2000.

A. Odena, V. Dumoulin, and C. Olah. Deconvolution and checkerboard artifacts. *Distill*, 1(10):e3, 2016.

J. Pennington, R. Socher, and C. Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.

I. Poggi and F. D'Errico. Cognitive modelling of human social signals. In *SSPW@ MM*, pages 21–26, 2010.

R. Ratcliff and P. L. Smith. A comparison of sequential sampling models for two-choice reaction time. *Psychological review*, 111(2):333, 2004.

N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.

P. Schultes, V. Dorner, and F. Lehner. Leave a comment! an in-depth analysis of user comments on youtube. *Wirtschaftsinformatik*, 42:659–673, 2013.

A. Severyn, O. Uryupina, B. Plank, A. Moschitti, and K. Filippova. Opinion mining on youtube. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, page 1252–126, 2014.

R. C Shiller. Irrational exuberance. *Philosophy and Public Policy Quarterly*, 20(1):18–23, 2000.

S. Siersdorfer, S. Chelaru, W. Nejdl, and J. San Pedro. How useful are your comments?: analyzing and predicting youtube comments and comment ratings. In *Proceedings of the 19th international conference on World wide web*, pages 891–900. ACM, 2010.

H. A. Simon. A behavioral model of rational choice. *The quarterly journal of economics*, 69(1):99–118, 1955.

C. Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

C. Sims. Rational inattention and monetary economics. *Handbook of Monetary Economics*, 3:155–181, 2010.

N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *Journal of machine learning research*, 15(1):1929–1958, 2014.

I. Sutskever, O. Vinyals, and Q. Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.

A. Tutino. 'rational inattention' guides overloaded brains, helps economists understand market behavior. *Economic Letter*, 6, 2011.

A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(4481): 453–458, 1981.

L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov): 2579–2605, 2008.

T. Van Erven and P. Harremos. Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014.

V. Vapnik and Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. In *Measures of complexity*, pages 11–30. Springer, 2015.

H. R. Varian. The nonparametric approach to demand analysis. *Econometrica: Journal of the Econometric Society*, pages 945–973, 1982.

H. R. Varian. Nonparametric tests of models of investor behavior. *Journal of Financial and Quantitative Analysis*, 18(3):269–278, 1983.

H. R. Varian. Non-parametric analysis of optimizing behavior with measurement error. *Journal of Econometrics*, 30(1-2):445–458, 1985.

H. R. Varian. Revealed preference and its applications. *The Economic Journal*, 122(560):332–338, 2012.

A. Vedaldi and K. Lenc. Matconvnet: Convolutional neural networks for matlab. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 689–692. ACM, 2015.

P. Vincent, H. Larochelle, Y. Bengio, and P. Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pages 1096–1103. ACM, 2008.

C. Wirth, J. Furnkranz, and G. Neumann. Model-free preference-based reinforcement learning. In *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, pages 2222–2228, 2016.

C. Wirth, R. Akrour, G. Neumann, and J. Fürnkranz. A survey of preference-based reinforcement learning methods. *The Journal of Machine Learning Research*, 18(1):4945–4990, 2017.

M. Woodford. Inattentive valuation and reference-dependent choice. *Unpublished Manuscript, Columbia University*, 2012.

J. Xie, R. Girshick, and A. Farhadi. Unsupervised deep embedding for clustering analysis. In *International Conference on Machine Learning*, pages 478–487, 2016.

D. Xu and D. Erdogmuns. Renyi's entropy, divergence and their nonparametric estimators. In *Information Theoretic Learning*, pages 47–102. Springer, 2010.

A. L. Yarbus. *Eye movements and vision*. Springer, 2013.