Category learning is the inference of category structures from a set of stimuli each labeled as belonging to one of the categories. Human category learning is of central importance to many ideas in psychology. One of the most basic traits of human cognition is the ability to group objects in the world into cohesive categories. Despite the apparent simplicity of this problem, people display quite complex behaviors in categorization settings. Their high degree of insightfulness and intelligence has been verified both anecdotally and through laboratory experiments. Some features of human-level categorization are: selectively attending to salient features, generalizing categories to novel objects, forming new categories to explain surprising data, generating hierarchical category structures, estimating confidence by deciding how loosely to generalize, and transferring knowledge between tasks about what types of objects tend to be categorized together.

Classical models of human category learning have not been able to explain all of these phenomena. The hierarchical Dirichlet process (HDP) is a framework that can be used to specify rich models of human categorization, both subsuming many existing models and containing promising new ones as well.

Psychological models of categorization fall into three general categories: Prototype models (Reed, 1972) represent a category by a single object, the prototype. The strength of a new stimulus’ membership in the category is measured by its similarity to the prototype. Exemplar models (Medin & Schaffer, 1979) represent a category by memorizing every instance of it, the exemplars. The strength of a new stimulus’ membership in the category is measured by its average similarity to the exemplars. Intermediate models (Anderson, 1990; Rosseel 2002; Vanpaemel et al., 2005) represent a category by clustering its instances and computing the strength of a new stimulus’ membership in the category by its average similarity to the cluster centers. This is equivalent to a prototype model when only a single cluster is used, and to an exemplar model when every object is in its own cluster. The most interesting cases are between these extremes.

In theory, intermediate models must consider all partitions of a category’s instances. In previous work, this limitation has been skirted by using suboptimal, greedy clustering algorithms and/or assuming that the number of clusters is fixed ahead of time. Due to efficient sampling algorithms such as Markov chain Monte Carlo (MCMC), more sophisticated models based on Bayesian inference, such as the HDP, can now be tested.

Introducing the Hierarchical Dirichlet Process

The Hierarchical Dirichlet Process (HDP) is a statistical model that can be used to represent the way in which humans learn categories. The HDP is a non-parametric Bayesian model that allows for the hierarchical organization of data into clusters at different levels of abstraction. This allows for the representation of categories with varying numbers of subcategories.

The HDP is defined as a distribution over distributions, where each distribution is a Dirichlet distribution. The parameters of the Dirichlet distributions are themselves drawn from other Dirichlet distributions, creating a hierarchical structure. This hierarchical structure allows the model to automatically determine the number of categories and subcategories, as well as their relationships to each other.

The HDP model can be used to represent the way in which humans learn categories. For example, the model can be used to represent the way in which humans learn to categorize objects based on their shape and color. The model can automatically determine the number of shape and color categories, as well as the relationships between them.

The HDP model has been shown to be effective in modeling human category learning tasks, such as the Archers task and the Car and Cat task. In these tasks, participants are presented with a set of objects and are asked to categorize them into different categories. The HDP model is able to accurately predict the categorization responses of human participants.

Conclusion

The Hierarchical Dirichlet Process is a powerful tool for modeling human category learning. It allows for the automatic determination of the number of categories and subcategories, as well as the relationships between them. The HDP model has been shown to be effective in modeling a variety of category learning tasks, and has the potential to be applied to a wide range of other cognitive tasks as well.