

# クラウドソーシングを用いた 半教師あり学習のための深層生成モデル

## 背景・問題設定

比較的低コストでラベル付きデータを集める方法としてクラウドソーシングが注目

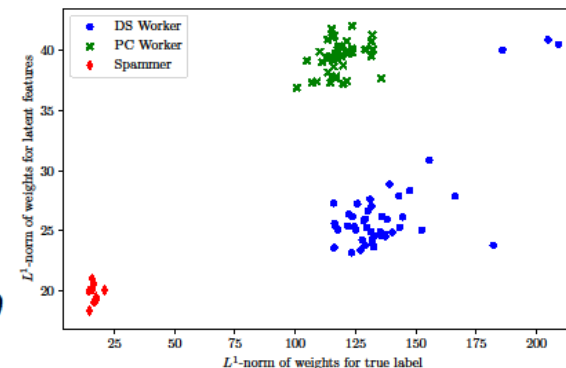
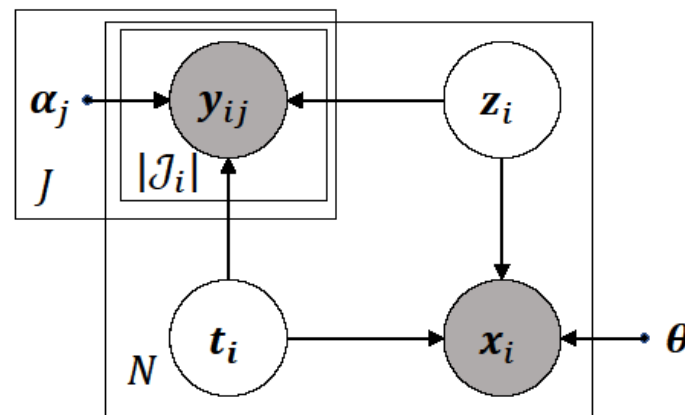
- 😞ワーカーの作業品質にはばらつきあり
- 😞得られる全てのデータにラベルを付けることは非現実的

→クラウドソーシングで得られたラベル付きデータ + ラベルなしデータからの学習

## 提案手法

ワーカーのラベリング過程に関する生成モデルに基づいた方法を提案

- 真のラベルと潜在特徴を潜在変数として導入
- データの特徴ベクトルに関する分布も導入し、深層ニューラルネットを用いて表現



実データ・人工データを用いた実験で既存手法を上回る性能  
ワーカーのラベリングの特性を正しく学習