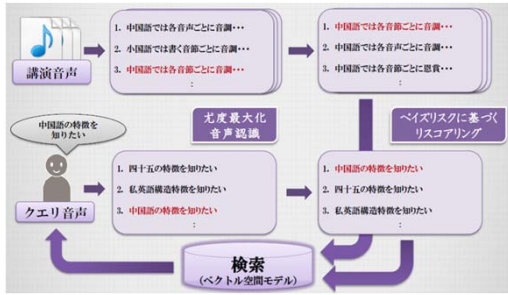


音声処理技術と高齢者・障害者支援への応用

情報学領域 情報科学系列 准教授 西田昌史

音声認識と話者認識

音声入力による音声ドキュメント検索



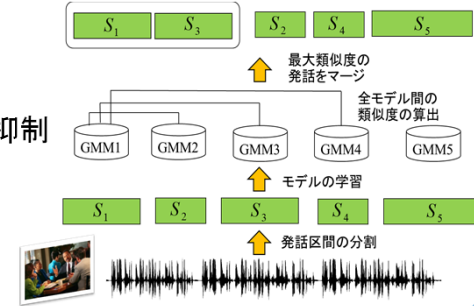
多人数会話における発話の自動書き起こし



咽喉マイクによる雑音・発話重畳の抑制



多人数会話における発話の自動分類



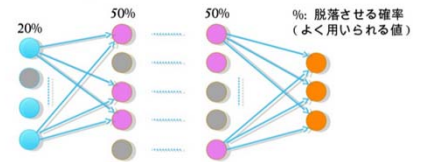
高齢者の見守りシステム

スマートフォンで収集した音や加速度情報から日常生活行動を認識



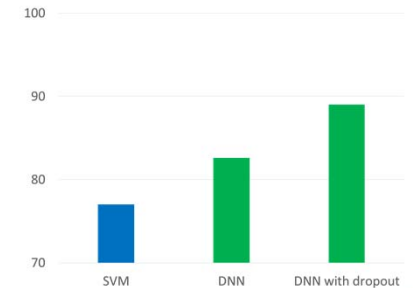
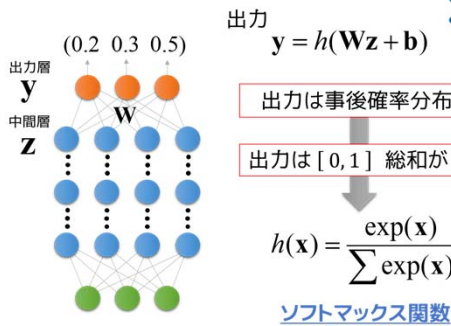
72時間連続で屋内外における日常生活行動データベースを構築 (合計1,400時間)

Dropout [Hinton+, 2012]



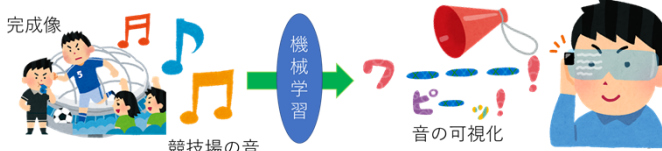
- Deep Learning に関する最も重要な技術の一つ
- 学習時、毎回ランダムに選んだ入力・隠れユニットを一旦取り除いて学習する (サガらせる)
- ユニット間、パラメータ間の依存関係が減る
- アンサンブル学習の近似になる

DNNで多クラス分類



聴覚障害者への支援

ARグラスによる競技場の音の可視化



話者A: じゃ城壁のある街の左要するに
話者B: 上行く左上行く
話者A: 左上の行つてでそれその絵の左側にそつて真下に行けばいいの
話者B: そうそうそう 真下にうん
話者A: 真下に行くとき度ゴーストタウンであるでしょう
話者B: ゴーストタウン?
話者A: そつちにはないのか
話者B: ないな

テキストへの発話印象の可視化

話者A: じゃ城壁のある街の左要するに
話者B: 上行く左上行く
話者A: 左上の行つてでそれその絵の左側にそつて真下に行けばいいの
話者B: そうそうそう 真下にうん
話者A: 真下に行くとき度ゴーストタウンであるでしょう
話者B: ゴーストタウン?
話者A: そつちにはないのか
話者B: ないな

Kinectと3Dモデルを用いた手話学習支援システム



keyword: 音声処理, 音声認識, 話者認識, 行動認識, 機械学習, 高齢者, 障害者