

# 実アプリケーションで機械学習をうまく活用するための10の心得

Google 株式会社

工藤 拓

taku@google.com

自然言語処理において機械学習を積極的に取り入れている研究やアプリケーションが増えてきている。年次大会の発表リストを見ても、機械学習を用いていない研究を探すほうが困難なほど、機械学習は有効な道具としてのその地位を確立しつつある。研究者の多くは機械学習の持つセクシーな性質に魅せられ、自身の研究に情熱を傾け、実際に多くのタスクにて成功を収めている。今後この流れはさらに加速していくであろう。

しかしながら、機械学習を実アプリケーションに応用し、うまく活用することは、思っているほど簡単なことではない。研究段階では、「精度」という単一の尺度で手法の良し悪しが評価されるが、実応用においては、「精度」は数ある評価基準の1つのすぎない。他の評価基準によっては、機械学習を用いることが困難になったり、既存のルールベース手法が高く評価されることもある。機械学習と言語処理をメインに研究してきた(発表者のような)研究者にはなんとも耳の痛い話ではあるが、そんな時に地団駄を踏んで感情的になったり、ことの成り行きを静観するのではなく、客観的な事実を積み重ね、機械学習の有効性をアピールしていく必要がある。

本発表では、「実アプリケーションで機械学習をうまく活用するための10の心得」と題し、あなたの言語処理アプリケーションを成功させるための具体的でかつ実践的なノウハウを紹介する。具体的には、発表者の経験を踏まえ、どのような方法論やモデルが実アプリケーションにおいて生き残るか、淘汰されるか、具体例を交え紹介する。本発表が開発者の実践的な指南書にとどまることなく、言語処理研究の新しい応用や方向性を示す起爆剤になれば幸いである。