

審査の結果の要旨

氏 名 全 弘 宇

本論文は、医療応用上きわめて重要な病気と遺伝子の関係を文献からマイニングする手法について論じたものである。本論文では、病名、遺伝子名の抽出を最大エントロピー法を用いて行い、さらに関係の抽出の際にそのカテゴリーの予測も行う手法を提案している。さらに生物医学論文要旨データベースから前立腺がんと胃がんに関係する遺伝子について実験を行い、その効力を実験で検証している。

本論文の第一章では、本研究が必要とされている背景について、生物医学論文要旨データベースMedlineの文献数の増大等の観点から概観している。

第二章は本論文に関連する従来研究として、文献からタンパク質-タンパク質関係の抽出する研究と病気-遺伝子関係の抽出する研究を挙げている。また、第三章では、本論文の提案する手法のおおまかな流れを説明している。

第四章では、文献から病名と遺伝子名を抽出する前段階で用いる病名と遺伝子名の辞書の構築について、および、生物学者によって注釈を与えられた病気-遺伝子の関係についてのコーパスの作成について述べている。

第五章では、前章の辞書のマッチングによって文献から得られた病名および遺伝子名について、それらが実際に関連しているかどうかを最大エントロピー法を用いて判定する手法について提案し、その性能を実験を通して検証を行っている。

さらに第六章では、こうして得られた病名と遺伝子名の候補を、病因としての関係と臨床的指標としての関係に分類する最大エントロピー法に基づく手法を提案し、その性能を実験を通して検証している。また、論文末には補遺として、実験で誤って予測された病気-遺伝子関係の実例を多数あげ、このシステムを実際使用する際の参考となるようにしている。

第七章では、提案した手法の性能を多数の実験を通して検証している。また、メモリベース学習法を適用した場合よりも本提案手法によるものの方が性能がよいことを示している。

第八章では、以上の成果を総括し、さらに今後取り組むべき問題についての展望をしめしている。

以上のとおり、本論文は医療応用上きわめて重要な病気と遺伝子の関係を文献からマイニングする新しい手法について述べたものであり、従来手法との比較からもその優位性は明らかである。なお、本論文は、論文提出者が主体となって分析及び検証を行ったもので、論文提出者の寄与が十分であると判断する。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。