

情報工学専攻 2年  
徐 研究室

赤木 里騎 さん

第71回 電気・情報関係学会九州支部連合大会  
『奨励賞』 受賞



一般社団法人 電気・情報関係学会九州支部連合大会委員会主催の平成 30 年度（第 71 回）電気・情報関係学会九州支部連合大会が去る平成 30 年 9 月 27 日・28 日に大分大学 巨野原キャンパスにて開催され、本学工学研究科 情報工学専攻 赤木里騎さんの発表が、『奨励賞』を受賞しました。

この賞は大会において優れた研究発表を行った若手研究者を表彰するもので、約 420 件の発表に対しわずか表彰 6 件の狭き門です。

この大会で受賞した研究テーマは「分散表現モデルにコーパスの差異が与える影響について」です。

情報工学専攻 2年 赤木 里騎さん（鹿児島県 大口高校出身）

【 分散表現モデルにコーパスの差異が与える影響について 】

概要： 自然言語処理の分野では Mikolov らによって提案された Word2vec が構築する、単語をベクトル化した分散表現を用いた研究が増えている。分散表現は単語同士の類似度を算出し、単語と単語の演算をすることを可能にしている。

本学では内定報告書を Web 上で検索・閲覧できるシステムを運用している。検索は企業名や業種・職種などの検索が可能であったが、年々内定報告書は増え、利用者に役立つ内定報告書の推薦が困難な状況となっていた。そこで内定報告書に記述される「後輩へのアドバイス」のような自然言語情報を活用して、利用者が検索時に入力した単語と報告書に記述された文章の分散表現から類似性を測り内定報告書を推薦するシステムの開発を進めている。

分散表現の学習にはコーパスが大きいほど良いとされるが、分散表現モデルを適用するシステムとの関連性も重要である。本研究では分散表現モデルの比較のために 3 つ用意した。内定報告書と本学の学生が記述したキャリアポートフォリオのデータに加えて、Wikipedia や Twitter から収集した大規模なテキストからコーパスを構築した。3 つのコーパスから作成した分散表現モデルを使用して推薦精度を評価し、コーパスサイズは Wikipedia や Twitter に劣る内定報告書と本学の学生が記述したキャリアポートフォリオのデータから学習した分散表現による類似度計算が有効であることを示した。



掲示期間 H-30-254  
1月31日～2月20日

この件のお問い合わせは広報課へ

徐 教授 と