

卒業論文

計算問題の特徴分布に基づく類題選出による自
己学習支援

Self Learning Support by Automatic Selection of
Calculation Exercises based on Feature Distribution of
Exercises

成蹊大学理工学部情報科学科

S152114 宮地 雄也

要旨

ポイント

序論と結論の内容をもとに研究の内容をまとめる.

- 問いは何か??
- 主張は何か??
- 結果はどうだったのか?
- 得られた成果の意義は?

目次

第 1 章	序論	1
1.1	memo	2
第 2 章	背景知識	3
2.1	アダプティブラーニング	3
2.2	分散表現	3
2.3	LSTM(Long short-term memory)	4
2.4	Attention	4
第 3 章	提案手法（章題は変える）	5
3.1	システム全体の流れ	5
3.2	計算式の特徴量抽出	5

第 1 章

序論

ポイント

問題提起を行う．解く価値があり，簡単には解けず，誰も解いていない問題を扱っていることがわかるようにする．

- どういう問題に取り組んだのか？
- その問題を解くことがなぜ重要なのか？ 社会的意義（有用性）・学術的意義（問題の面白さ）
- その問題はどこが難しいのか？ なぜこれまで解かれていなかったのか？ これまではどうしていたのか？
- その問題をどのようなアプローチで解こうとしたのか？ なぜそうしたのか？

昨今，小・中学生の理系離れが問題視されている．平成 30 年度全国学力・学習状況調査（全国学力テスト）の結果では平均正答率は小学校では算数 B が 51.7%，中学校数学では 47.6% とどちらも最も低く，ついで国語，理科の順で正答率が低く，結果として理系教科の習熟度が低いことを示している．この要因の一つに，数学は一つの計算方法が様々な分野に横断していくことが一度，苦手を生んでしまったらそこからの分野の理解度が下がり，次の分野での応用がきかないために連鎖的に苦手を増幅するように思う．各単元のちょっとした積み残しが，後々，尾を引いていることが全国学力テストの結果から見て取れる．この状況を打破するには子供一人一人の苦手と向き合い，苦手と感じる前に理解していくしかない．しかしながら，生徒と向き合うべき教師の労働時間は，... <悪さについてノベル>

この打開策として，Edtech（エドテック）に期待が集まっている，<いいことつらつら>その中でも，生徒一人一人に対し，最適な学習を提供する Adaptive Learning(アダプティブラーニング)に可能性を感じた．これを数学教育で効果的に用いれば，つまり子供を減らせるのではないかと考えた．しかし，教育の情報は，生徒の情報と結びついているためオープン化できないという問題がある．また数学の苦手に対してどのような手法を用いれば効果的なのかも未だ模索中である．

そこで計算式自体の特徴を抽出し，間違えた問題と同様の特徴を持つ問題が復習する類題として最適なのではないかという仮定のもと，本論文では数式の特徴を掴むために自然言語処理の分野で使われる分散表現を適用し，さらに再起ニューラルネットワークを用いて数式ベクトルを作り出すことを目標とし，そのベクトルを用いて実際に復習問題生成を行った．

1.1 memo

近年，子供達一人一人に目を向けたアダプティブラーニングが注目を集めていますが，現在の日本の教育現場では集合教育が基本であり一人一人に対応するのは難しい．また教員の働き方にも是正が求められていて，仮に個々に対応しようとする労働時間外になってしまう．そこで個人最適化した学習を情報技術で叶えようとする動きが活発である．そこで私は間違えた問いに対して復習問題を出す際，その数式の特徴を捉えて類題が作れないか考えた．今までは問題をといた子供達の集合に目を向けられていたが，本論文では数式自体に目を向け類似する問題を作り出そうとした．本論文では数式の特徴を掴むために自然言語処理の分野で使用される分散表現を適用し，さらに再起ニューラルネットワークを用いて数式ベクトルを作り出すことを目標とし，そのベクトルを用いて実際に復習問題生成を行った．

第 2 章

背景知識

ポイント

以降の内容を理解するための準備を行う。

- 章題は適切なものに変えること。章をわけてもよい。
- 以降の説明で用いる専門用語・表記法を説明する。
- 以降の内容を理解するのに必要となる、技術や理論を説明する。

2.1 アダプティブラーニング

今世の中でどのようなことが行われているかを収集。その手法を 2, 3 事例あげて説明

アダプティブラーニングとは、学習者一人ひとりの学習進捗度 (学習進度) に最適化された学習方法と学習教材を選択し、提供する仕組みを持つ学習エンジンやシステム、ソフトウェア、サービスなどを統括的に指す言葉です。

アダプティブラーニングは、IT 技術を教育分野に活用する EdTech(Education Technology) の 1 つとして世界中から多くの注目を集めています。アダプティブラーニングの先駆企業であり世界中でシェアを獲得している米 Knewton 社 CEO のライアン・プリチャード氏は、日本法人であるニュートンジャパン株式会社主催のアダプティブラーニングをテーマにしたイベント『Knewton Day Tokyo 2017- Adaptive Learning Summit -』の Q & A セッション内で、アダプティブラーニングの定義について上記のように回答しました。よって本論文中でも以下のように定義する。

データをもとにパーソナライズされた経験を継続的に提供するもので、生徒がシステムを使うたびにコンテンツやモデルをアップデートし、最適な道筋をアップデートするもの

日本国内での様子..

2.2 分散表現

様々な手法を紹介する CBow, skipGram, GloVe など計算式を並べて最後にどこが長所で、どこが違うのかを表現

2.3 LSTM(Long short-term memory)

通常の RNN では叶わないところを明確に

2.4 Attention

これからね

第 3 章

提案手法（章題は変える）

ポイント

自分の提案する解決方法を説明する.

- 章題は適切なものに変えること. 章をわけてもよい.
- 必ず具体例を用いること.
- 最初に問題を解く上で最も難しい点とそれを解決するアイデアを示す.
- 詳細については, 全体の流れを示した後, 各ステップについて説明する.
- 検討時に行った予備評価の結果があれば示す.

3.1 システム全体の流れ

<図をいれながら>

3.2 計算式の特徴量抽出

3.2.1 概要

<idea> (内容充実させる) 数式を分布化する際, そのベクトルの中に数式の特徴を入れ込んだベクトルを生成する手法が確立していない. そこで本論文では数式の各文字, 記号を単語のようにみなし, onehot ベクトルを作成し, それを埋め込み層で特徴を踏まえた低次元ベクトルに変換したのち, 系列変換モデルで読み込むことで低次元で数式の特徴を掴んだベクトルを生成できないかと考えた.

<手法> この考えを実現するために数式は我々が目にする $2x + 3 = 5$, $\frac{3x-1}{2} + 4 = \frac{2}{5}$ ではなく, テキスト化かつその特徴を強く受けた形に変換する必要がある. そこで本論文では数式をある一定のルールの中でテキスト化されている TeX 形式の数式を用いる. 上記の計算式なら `2x+3=5,\frac{3x-1}{2}+4=\frac{2}{5}` とし, このテキストデータを用いて文字単位の埋め込んだベクトルを作成する.

実験を行った手法は以下の三種法で行い, それぞれ分布を python を用いて確認した.

- CBOW
- SkipGram
- ...