

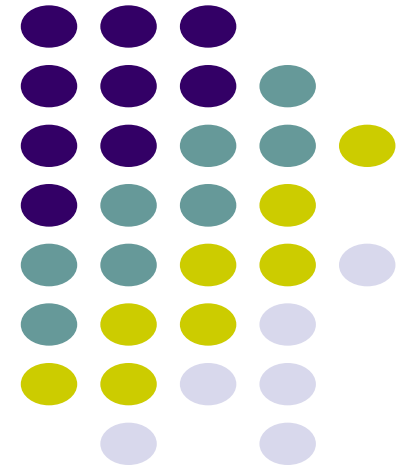
言語処理工学

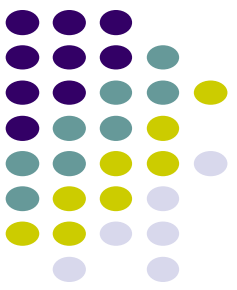
期末レポート課題

2025年度

提出期限：8月10日（日）

！レポート作成にあたっては、この資料を隅々までよく読むこと

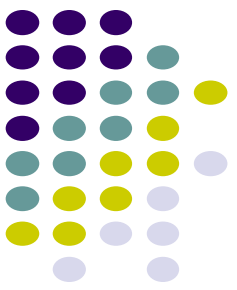




概要

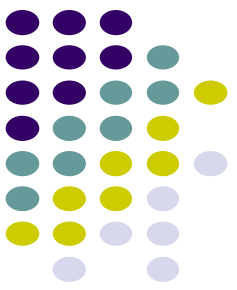
- 全3課題、全てについて取り組むこと
- 内容をレポートにまとめて Moodle から提出

※ 成績は中間レポート・期末レポート・講義出席(ない回は講義視聴)を合わせて総合評価します



レポートの要件

- 課題1, 3は指定枚数以上を書く
 - 図表をページ数に含めてよい
 - 異常に大きい図や大きいフォント、改行の多用など、意図的な水増しが見られるものは大幅減点対象とします
- 講義資料以外の資料を自分で調べて書いてください
 - 講義資料をまとめただけのものはNGです
 - 最低1つは講義資料以外の資料を使ってください
- 調べた参考資料の情報をすべてレポート末尾に列挙してください
 - タイトル・著者(執筆者)・文献情報 など
 - まとめ記事・ブログ記事、SNS投稿等は不可
 - 参照情報がない課題は大幅減点とします

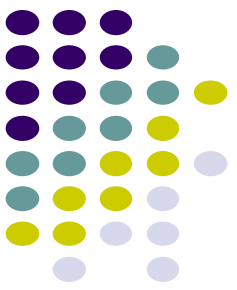


課題1

(1) 自然言語処理の手法は、大きく分けて文法や事例等のルールを用いる方法と、コーパスに基づく統計的な方法の2種に分類される。それぞれの方法について長所・短所をまとめ、具体例を挙げながら説明せよ。(1ページ)

(2) ニューラルネットワークモデルの構成要素と動作原理を説明し、自然言語処理における同モデルを用いた処理の例を1例、具体的に挙げて解説せよ。(2ページ)

(講義資料以外に自分で調べたことを書くとベター)



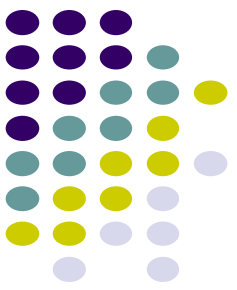
課題2

ある湖に、A～Fの6種類の魚が存在する。この湖で何回か魚の捕獲を行ったところ、以下のような結果が得られた。

1回目: A 3匹、B 6匹、C 9匹、D 0匹、E 2匹、F 0匹
2回目: A 2匹、B 15匹、C 12匹、D 0匹、E 0匹、F 1匹
合計: A 5匹、B 21匹、C 21匹、D 0匹、E 2匹、F 1匹

この池における魚A～Fの出現確率 $P(A) \sim P(F)$ を以下のそれぞれの方法で推定し、結果を比較せよ。

1. 最尤推定
2. 加算法 ($\delta=1$)
3. 削除推定法



課題2補足

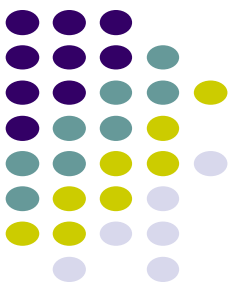
講義第8回の資料では削除推定法における補正出現回数を以下のように定義していたが、

$$\frac{C_r^{01} + C_r^{10}}{N_r^0 + N_r^1}$$

この式では $N_r = 0$ となる頻度 r があるとき確率の合計が1とならない。

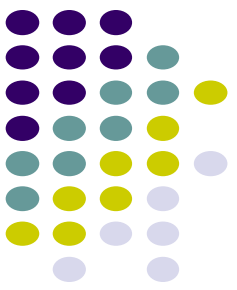
本レポートでは、代わりに以下の式を用いること。

$$\frac{C_r^{01}}{N_r^0} + \frac{C_r^{10}}{N_r^1}$$



【課題3】

- 以下のキーワードのなかから1つを選び、その言葉の意味や動作原理、応用事例について調べ、自分なりの見解を加えてまとめよ（3ページ以上）。
 - TRANSFORMER
 - 自動音声翻訳
 - マルチモーダルLLM
 - NaturalSpeech 3



提出方法

- レポート
 - 最初のページに「講義名」「氏名」「学籍番号」を必ず書くこと
 - PDF形式のみとする(.doc や .tex は受け付けません！)
 - 全課題の回答をまとめて1つのPDFにしてください。
- Moodle から提出