

知能情報実験 III（データマイニング班）
発表タイトルやテーマ名

グループの学籍番号列挙 xxx, yyy, zzz

提出日：20xx 年 x 月 x 日

目次

1	はじめに	0
1.1	ストーリーの読み取れない章立て	0
1.2	段落のない本文	0
1.3	出典なしのストーリー	1
1.4	箇条書きだけの本文	1
1.5	整合性の取れていない様式	1
1.6	実験の目的と達成目標	1
1.7	テーマ**とは	1
2	実験方法	2
2.1	実験目的	2
2.2	データセット構築	2
2.3	モデル選定	2
2.4	パラメータ調整	2
3	実験結果	2
4	考察	3
5	意図していた実験計画との違い	3
6	まとめ	3

概要

概要本文書は知能情報実験 III（データマイニング班）におけるレポートのテンプレートとして用意したものである。一般的な実験レポートに関する補足と共に、データマイニング班における実験レポートに求められる内容を確認するために用意した。ここに書いてある事柄は全てを必須とするわけではなく、適宜取捨選択や追加編集してもらって構わないが、実験報告書としての位置づけを忘れずに利用すること。

1 はじめに

~~~~~ここから削除対象~~~~~

この文章は提出時には全て削除すること。レポート作成時には以下の点に留意すること。またより詳細は [1] を一読することを強く勧める。

### 1.1 ストーリーの読み取れない章立て

原則として報告書の読み手の立場に立って可能な限り読みやすく章立てを検討すべきである。授業で作成指示された課題レポートについては必ずその担当教員が読むべきものであるが、より一般的に報告書を読む状況とは、そもそも自身にとって読む必要があるかを随時判断しながら読み進める、もしくは対象外と判断してスキップするものである。

より具体的には、タイトルを眺めて読む対象かどうかを考え、興味を持てたならば概要を読み、その上で詳細を知りたいと思ったら章や節のタイトルを眺めて希望する箇所を選んで読み、更に全体を把握したい場合によく冒頭から全てを読み進めることになる。別の観点から述べると、冒頭から全ての本文を読み進め、最後まで読み通してようやく「これは読むべき報告書ではなかった」と気付かされるような書き方になっている報告書は、書き方がまずい。

このような状況に陥らないようにするため、いきなり本文を書き始めるのではなく、箇条書きやアウトライン機能のあるエディタを使い章立てを検討すべきである。

### 1.2 段落のない本文

既にある章や節を読むかどうかをタイトルから判断することを述べた。この次に検討する単位が段落（パラグラフ）である。あまりにも長すぎる本文は読みにくく、結果として読むべきか否か判断しづらい。このため内容のまとまりを考えて段落に分けて書くことを意識しよう。

### 1.3 出典なしのストーリー

卒業研究ではないためそこまで重視しないが、可能な範囲で出典を明記しよう。例えば「スパムメールは増える傾向が続いており、この検出はとても重要だ。」というような背景を書こうと思った場合、そもそも本当に増加傾向なのか、そのことを述べている調査報告書等はないのか。もし見つからないにも関わらず書いてしまうのであれば、それは一種の捏造である。今回の学生実験ではそこまで求めないが、可能な範囲で出典探しをしよう。

### 1.4 箇条書きだけの本文

本文なしに箇条書きだけで終わっている章や節は、そもそも書き方が不適切である。α版やβ版を検討中に箇条書しておくことで全体構成を検討する状況なら良いが、最終版で箇条書きだけとなるのは読み手からする何故これらを列挙したのか意図を汲み取れず、報告書としての体をなしていない。列挙する前に十分に本文で説明しよう。

### 1.5 整合性の取れていない様式

例えば句読点について「、。」や「,.」はたまた「,。」等が混ざって書くことは避けよう。段落替えについても全体を通して統一した書き方にしよう。「です・ます調」「である調」についても統一しているならばどちらであっても構わないが、「である」の方が望ましい。この理由については [1] の pp.12-14 を参照すること。

~~~~~ここまで削除対象~~~~~

1.6 実験の目的と達成目標

知能情報実験 III は、情報工学分野のより専門的な知識を理解・習得することを目的として、半年間でシステムの開発やデータ解析等に取り組む実施される。その中の一つデータマイニング班においては機械学習外観ならびにその応用を通し、対象問題への理解、特徴量抽出等の前処理、バージョン管理やデバッグ・テスト等を含む仕様が定まっていない状況下における開発方法、コード解説や実験再現のためのドキュメント作成等の習得を目指す。

1.7 テーマ**とは

選んだテーマについてどのようなテーマなのか、そのテーマをやる意義について述べよう。200字以上。

(例) 本グループでは**における**することを対象問題として設定した。**とは [2] によると~~~であり、**を**することで**に寄与する。また [3] によると~~~とも述べられており、

2 実験方法

実験手順を過去形で述べよう。日誌のように時系列ではなく、成果物として報告する最終版を再現するための実験手順が良い。第三者が再現するために必要な手順であることが重要だ。また、列挙した項目毎に具体的な内容を subsection で述べよう。

(あくまでも例です)

1. 実験目的
2. 実験計画
3. データセット構築
4. モデルの選定
5. パラメータ調整

2.1 実験目的

実験を通して明らかにしたいこと、確認したいこと、検証したいことを述べよう。

2.2 データセット構築

既にどこかで公開されているデータセットをダウンロードして利用したのならば、その URL を掲載する程度で構いません。独自構築した場合にはその構築方法を述べよう。

2.3 モデル選定

どのようなモデルやアルゴリズムを利用したのか、何故それを選んだのか述べよう。

2.4 パラメータ調整

手動調整が必要なパラメータについて、どのように調整したのか述べよう。

3 実験結果

事実として得られた結果を示そう。なお、以下の点に留意すること。

- 「思う」「思われる」のような主観ではなく、客観的事実を述べること。
- 図表には適切なキャプションを付けること。
- 挿入した図表について、本文中でその読み方を述べること。その際には label, ref により相互参照すること。

- レポートにおけるグラフの作成においては、以下の点に注意する。
 - － 軸目盛および軸ラベルに関する注意事項
 - * 必ず軸ラベルを表示する
 - * 軸に単位がある場合には、ラベルに単位を付記する
 - * 軸目盛は適切な感覚で表示する
 - * 軸目盛は述べたい内容に応じて線形スケールと log スケールを使い分ける
 - * 印刷時に明瞭に読むことができるサイズで表示する
 - － 線・点・ポイントおよび凡例に関する注意事項
 - * 線・点・ポイントは、印刷時に明瞭に識別できる太さやサイズで表示する
 - * 1つのグラフに複数にデータを表示する際には、データごとに異なる線種、線の太さ、ポイント形状などを使用する
 - * モノクロ印刷でも識別できるように線・点・ポイントを使用することが望ましい
 - * 凡例は線・点・ポイントに重ならないように注意する

4 考察

実験課題への取り組みを通し、実験の意義、実験からわかったこと、今後の展望などを述べる。失敗やつまづきがあれば、それらについての失敗分析を含めると良い。

5 意図していた実験計画との違い

グループワークとして2ヶ月程度の時間が用意されていた。ガントチャート??等、何かしら工夫して全体の計画を述べよう。これらの期間をどのように使おうとし、実際どうだったのかについて自己評価（振り返り）してみよう。大きなズレがある場合それは何故起きたのか、どうやればそのギャップを縮められそうか検討してみよう。

6 まとめ

データマイニング班の達成目標を振り返り、選んだテーマに対する機械学習の適用を通して得られた知見や学んだことをまとめよう。また今後やるべきことや後進に伝えたいこと等あれば自由に述べよう。

参考文献

- [1] レポート作成の手引き レポートの基本的形式に関するガイド, <https://www.kanazawa-u.ac.jp/wp-content/uploads/2015/01/tebiki2.pdf>, 2020/07/02.
- [2] テーマ出典, 書籍 or 特定の記事 or webpage, webpage の場合は参照日も記そう, 2020/07/02.

[3] テーマ出典 2, 出典は半角,. で書こう.

[4] ガントチャート, <https://ja.wikipedia.org/wiki/ガントチャート>, 2020/07/02.