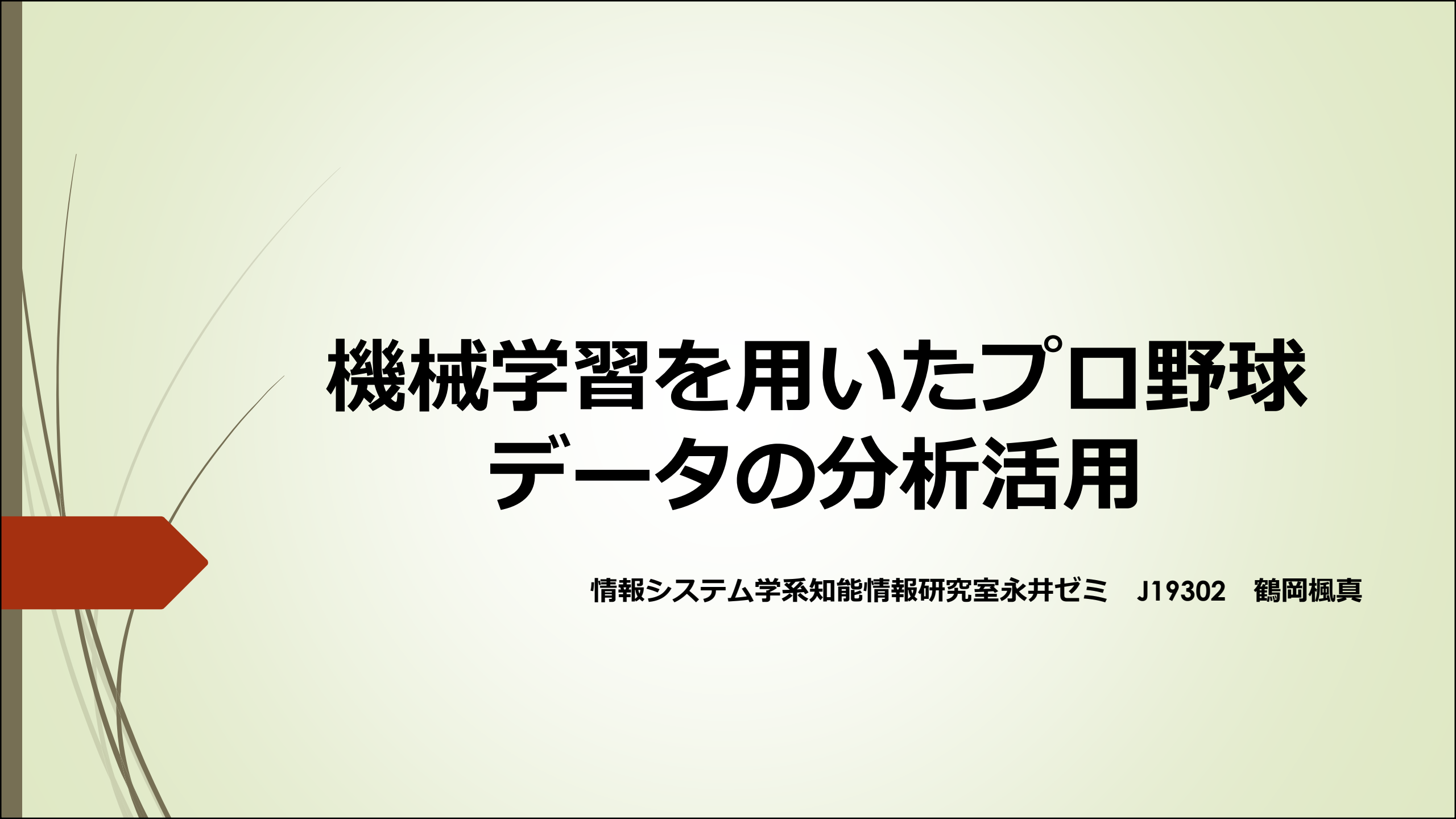




機械学習を用いたプロ野球 データの分析活用

情報システム学系知能情報研究室永井ゼミ J19302 鶴岡楓真



目次

1. 研究の動機と目的
2. 開発環境及び開発方法
3. 進捗状況
4. 今後の展開

研究の動機と目的

B9やGG賞は記者投票で選出される



近年、適当orただの記者の中の人気投票になっている



機械学習を利用して選ばれやすい選手の特徴を抽出



真の表彰選手を選出する

選ばれやすい選手の特徴の抽出

■ 開発環境

: Anaconda (jupyter notebook)

■ 方法

: 特徴量エンジニアリングで特徴を抽出

特徴プログラミングから予想される結果

- 打者のほうでは走力面の成績より打撃面の成績が良い選手の方が選出されている
- 守備のほうではエラーの個数や刺殺の個数が影響してくる

真の表彰選手の選出するシステムの提案

■ 開発環境

→ Eclipse

■ 方法

： 一定試合出場している選手の成績を入力

↓
そのシーズンの平均と比べ点数をつける

↓
上記で出した点数を比べ一番高い選手を選出

進捗状況

◆あまり進んでいない

- ・ 選ばれやすい選手の特徴を抽出
→ 特徴量エンジニアリングの勉強中
- ・ 真の表彰選手を選出
→ 簡易なものではあるがプログラム完成

名前	球団名	打率	本塁打	打点
B・レアード	ロ	0.262	29	95
鈴木大地	楽	0.277	10	53
T-岡田	オ	0.241	17	63
中村晃	ソ	0.245	8	56
高濱祐仁	日	0.262	8	43
山川穂高	西	0.232	24	66
平均		0.253	16	63



スコア
54
8
-11
-23
-19
-10

今後の展開

- 特徴プログラミング

- プログラムの作成及び修正

- システムの提案

- プログラムの修正及び評価方法の見直し

参考文献

- スポーツデータを用いた特徴量エンジニアリングと野球選手の成績予測

<https://shinyorke.hatenablog.com/>

- Python機械学習「ホームラン数予測」

<https://alivetodayblog.com/>

- スパイア【SPAIA】

<https://spaia.jp/>