

「わくわく DS 研究会」2025 年 9 月 会合録

日時：2025 年 9 月 20 日（土）10:00～12:00

場所／方法：Zoom による遠隔会議

出席者：協阪、下里、鈴木

1. 教育・学習カリキュラムに関する議論

- **学生向け勉強会の検討**：協阪氏より、学生向けにコンピュータの基本操作から段階的に学ぶ勉強会の開催について提案があった。
- **データサイエンス基礎講座（GCI）の共有**：下里氏より、東京大学の「データサイエンス基礎講座（GCI）」について詳細な説明が行われた。
 - 講座は全 15 回（約 150 時間）、Python を用いた分析や機械学習の基礎から実践的なコンペまでを含む。
 - 2024 年冬の講座には約 2 万人の応募があったが、原則として学生が対象である。
 - 社会人は「Omni Campus（オムニキャンパス）」を通じた申し込みが可能であり、9 月末が期限であることが共有された。

2. 発表：データサイエンスの最終課題と実務応用（発表者：下里）

- **IT 企業の離職率予測**：下里氏が受講したコースの最終課題として、「IT 企業の離職率予測と対策」について発表した。
 - 1,471 人の従業員データを LightGBM モデルで分析し、特徴量エンジニアリングを実施。
 - 分析結果に基づき、若手面接やインセンティブ提供などの対策を行うことで、効果的な離職抑制が期待できると報告された。
- **実務への応用（フィルム製造）**：勤務先におけるフィルム製造プロセスの改善に機械学習を活用しており、品質管理や環境条件の最適化に取り組んでいる事例が紹介された。
- **学習ツールと効率化**：Google Colab を使用した約 156 時間の学習過程において、Gemini Pro や GitHub Copilot を活用したバグ修正・コード生成が学習の助けになったとの報告があった。

3. 情報共有・討議

- **アルゴリズムコンテストへの参加**：鈴木氏より、画像検出アルゴリズムコンテストにエントリーしており、翌週に提出予定であることが報告された。
- **データサイエンスの役割**：協阪氏と下里氏の間で、従来の問題解決手法とデータサイエンスの違いについて議論が行われた。単純なグラフ分析では解決できない複雑な要因が絡む問題（「第三段階」の問題）に対して、データサイエンスが有効であるとの認識で一致した。
- **専門家雇用のコスト**：データサイエンスの専門家を雇用する場合のコスト（約 2,500 万円）と教育による内製化の効果について意見が交わされた。

4. 次回の例会

- **日時**：2025 年 10 月 25 日（土）10:00～12:00
- **発表予定者**：鈴木（テーマ：画像分類・物体検出に関する研究発表）

- 鈴木氏は予稿集の提出を控えており、コンテストの結果次第では11月以降に外部での口頭発表等の可能性がある旨が説明された。

以上