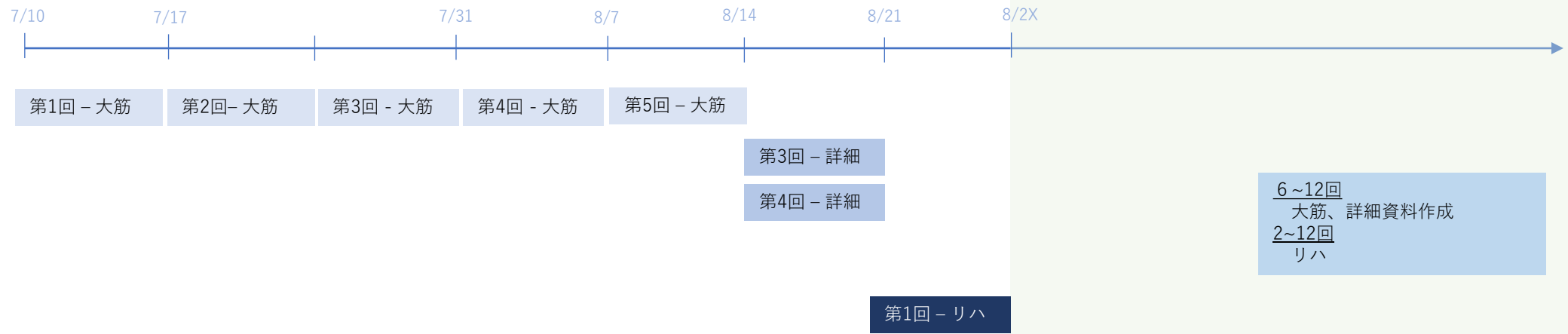


大筋資料作成（第1～5回まで）

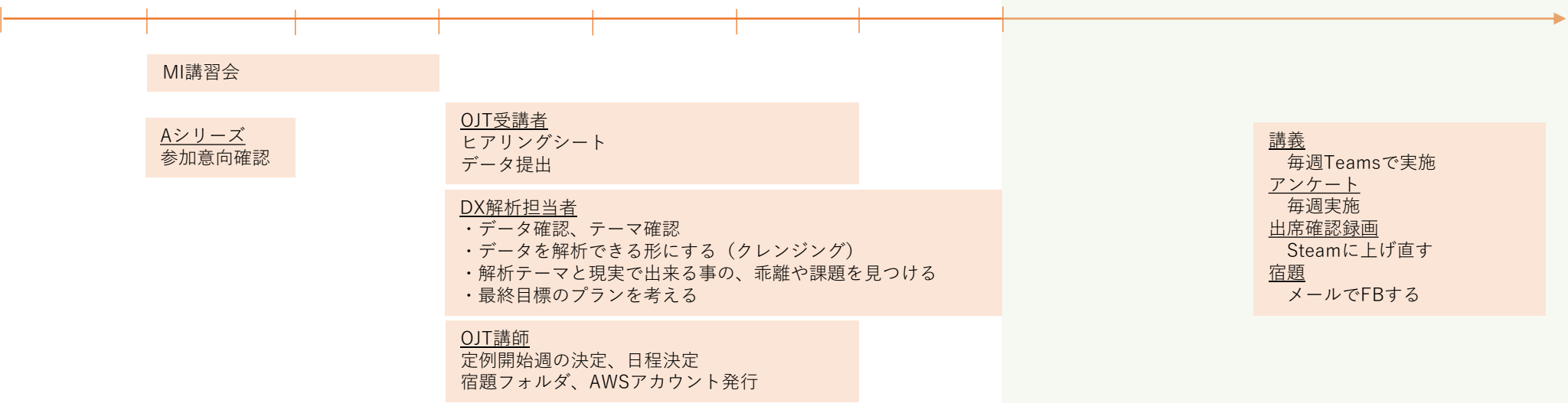
大筋資料作成（第6～12回）



7/24

OJT実施準備

OJT実施中作業



OJT講義のスケジュール

#	Before	After	
1	<ul style="list-style-type: none"> ✓ 解析OJTの位置付けと準備 (OJT内容、全体スケジュール) ✓ MI解析の手順(CRISP-DM) ✓ データセットの作成方法● ✓ MI解析ツール・AWS環境の説明● 	<ul style="list-style-type: none"> ✓ 解析OJTの位置付けと準備 (OJT内容、全体スケジュール) ✓ MI解析の手順(CRISP-DM) ✓ GUIツールの一連の流れを紹介 ✓ 解析設計書の作成方法 	<ul style="list-style-type: none"> OJT全体像の説明とGUIツールの簡単な紹介の時間としたい
2	<ul style="list-style-type: none"> ✓ 解析設計書の作成方法● ✓ 実験データの確認方法● 	<ul style="list-style-type: none"> ✓ GUIツールのStep1,2,3でのポイント説明(MI解析セミナー風) ✓ データセットの作成方法● 	
3	<ul style="list-style-type: none"> ✓ 特徴量エンジニアリングの行い方● ✓ 機械学習モデルの構築(回帰・分類・時系列)● 	<ul style="list-style-type: none"> ✓ 実験データの確認方法● ✓ 特徴量エンジニアリングの行い方● ✓ MI解析ツール・AWS環境の説明● 	<ul style="list-style-type: none"> データ加工技術をじっくり教える時間としたい 次回の為の準備
4	<ul style="list-style-type: none"> ✓ 機械学習モデルのオーバーフィッティングの考え方 ✓ 機械学習モデルにおける特徴量選択 ✓ モデル評価指標 ✓ モデル解釈(ステムプロット、SHAP値) 	<ul style="list-style-type: none"> ✓ 機械学習モデルの構築(回帰・分類・時系列)● ✓ 機械学習モデルのオーバーフィッティングの考え方 ✓ 機械学習モデルにおける特徴量選択 ✓ モデル評価指標 ✓ モデル解釈(ステムプロット、SHAP値) 	<ul style="list-style-type: none"> MLモデル構築の基本とモデルの解釈方法の基本を教える時間としたい 4回：特徴エンジのFB 5回：モデル解釈の情報共有
5			
6	<ul style="list-style-type: none"> ✓ モデル適用範囲、外れ値の考え方 ✓ 実験パラメータ探索の方法 ✓ 実験パラメータの選択 ✓ 実験結果の解釈のポイント 	<ul style="list-style-type: none"> ✓ 実験パラメータ探索の方法 ✓ 実験パラメータの選択 ✓ 実験結果の解釈のポイント ✓ モデル適用範囲、外れ値の考え方● 	<ul style="list-style-type: none"> パラメータ探索をじっくり教える時間としたい
7	✓ 機械学習モデルの見直し	✓ 機械学習モデルの見直し	<ul style="list-style-type: none"> これまでの取組を情報共有
8	<ul style="list-style-type: none"> ✓ 質疑応答、 ✓ 補足説明事項 ✓ 復習 	<ul style="list-style-type: none"> ✓ 質疑応答、 ✓ 補足説明事項 ✓ 復習 	
9	✓ 解析報告書の作成	✓ 解析報告書の作成	
10	✓ テーマ毎の報告・議論(1)	✓ テーマ毎の報告・議論(1)	
11	✓ テーマ毎の報告・議論(2)	✓ テーマ毎の報告・議論(2)	
12	✓ 全体まとめ、質疑応答	✓ 全体まとめ、質疑応答	

講義資料を分かり易く編集したい理由（目的）

講義内容が膨大である

- 12回のスライド数は400枚を超える
 - 講義の構成を再検討して、分かり易くしたい
 - スライドに優先順位を付けて、効率的に学習できるようにする

講義資料の表現が難解である

- 抽象的な表現や、専門的なスタイルの図が多い
 - 平易な表現に置き換える事で分かり易くする

実施内容自体が複雑である

- GUIツールの使い方を覚えるだけでも難しい
 - AWSにアクセスせず、GUIツールだけで解析できるよう簡潔にしたい
- 解析プロセス、理論、コツ、等が受講者にとっては新しい分野である
 - 優先順位を付けて講義を行う

講義資料作りのポイント

実践に必要な説明を優先

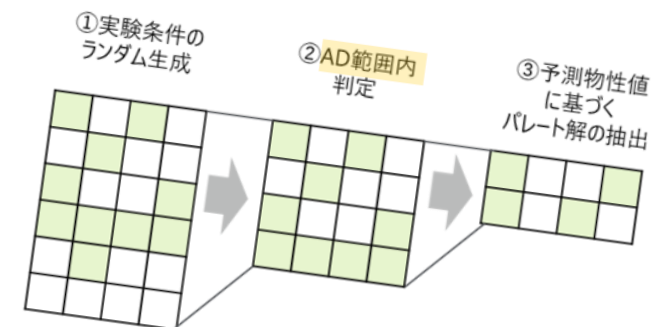
- 大筋を書いて、そこから詳細を派生させていく
 - GUIツールを動かす為に最低限必要な情報→解析を進める為に重要なポイント→詳細な理論

理論と実践が紐づき易い構成

- 説明している事は、GUIツールのどの部分の話か、常に分かる様に資料を描く

理解し易い表現

- 想像しにくい表現は簡易化する
 - 文字だらけの説明 → 想像できる図示にする
 - 細かすぎる説明 → 説明する上で必要な説明のみに削ぐ
 - 難解な図 → 別の例えを使う等で想像しやすくする



資料の組立て方（まず大筋 → 詳細を派生）

①データを入力して、GUIツールでStep3まで実行できる事

- 解析の大筋を理解すること
 - CRISP-DMの流れを理解できたか
 - GUIツールを使った解析の流れを理解できたか

OJT未受講でも満たせるレベル

②GUIツールに沿って解析を行う為の最低限な理論を理解できる事

- 具体的な理論
 - Step1：解析設計書を作成できたか
 - Step1：自身のデータを解析できる形式に変換する必要が理解できたか
 - Step1：データを確認する方法が理解できたか（散歩図、ヒートマップ、ヒストグラム）
 - Step2：機械学習モデルの特徴が理解できたか（過学習）
 - Step2：SHAP値の見方を理解できたか
 - Step2：特徴量エンジニアリングの基本を理解できたか（不要列削除、多重共線性、異なるモードのデータが混入、欠損値補完、正規化）
 - Step3：パラメータ探索の仕組みが理解できたか

OJT聞いたら満たせるレベル

③GUIツールを使って自身のテーマに②を適用できる事

- 具体的なアクション

-
-
-
-
-
-
-
-

OJT聞いて実践したら満たせるレベル

④GUIツールでは不要な、少し高度な理論を理解できる事

- 具体的な理論
 - LASSO正則化をロジックを理解する
 - 機械学習モデル作成の裏で起こっている事や、構造を理解する（様々な機械学習モデルの紹介、交差検証法の紹介）
 - 分子構造組み込みの活用方法を理解する

OJTの範疇を超えたアドバンスレベル

GUIツール：Step1,2,3それぞれの基本ポイント

講義初期段階で説明
MI解析セミナーみたいに
ダイジェストで説明する

Step1

- 解析設計書を作成方法
- 自身のデータを解析できる形式に変換する方法
- データを確認する方法
 - 散歩図、ヒートマップ、ヒストグラム…

Step2

- 機械学習モデルの特徴の理解
- SHAP値の見方の理解
- 特徴量エンジニアリングの基本理解
 - 不要列削除、多重共線性、異なるモードのデータが混入、欠損値補完、正規化

Step3

- パラメータ探索の仕組み理解