

地域社会 DX 推進パッケージ事業(AI 検証タイプ)

北海道函館市

観光地における連合学習を用いた多地点混雑予測
AI の学習データ多様化の実証

実証報告書

2026 年 1 月 30 日

AI01 函館市

ソフトバンク株式会社

国立大学法人東京大学、

宇宙サービスイノベーションラボ事業協同組合

目次

| | | |
|-----|-------------------------------------|----|
| 0. | エグゼクティブサマリ..... | 1 |
| 0.1 | 実証概要 | 1 |
| 0.2 | KPI/KGI の内容と達成状況..... | 1 |
| 0.3 | 考察..... | 2 |
| 0.4 | 成果..... | 2 |
| 0.5 | 課題..... | 3 |
| 1. | 実証の目的及び内容等..... | 4 |
| 1.1 | 実証の背景・取り組むべき課題..... | 4 |
| 1.2 | 実証の目的..... | 5 |
| 1.3 | 最終目標・構想イメージ..... | 6 |
| 1.4 | 本実証の位置づけ・目標..... | 7 |
| 1.5 | 社会実装に向けたロードマップ・取り組みの発展性..... | 8 |
| 2. | 業務実施体制..... | 10 |
| 2.1 | 実証機関..... | 10 |
| 2.2 | 実施体制図..... | 10 |
| 3. | 通信システム及び AI システムの構築..... | 12 |
| 3.1 | 通信システム及び AI システムの全体像..... | 12 |
| 3.2 | 通信システム及び AI システムの設置場所..... | 14 |
| 3.3 | 通信システム及び AI システムの構築・運用に要した費用..... | 21 |
| 3.4 | 通信システム及び AI システムの構築にあたっての留意事項等..... | 22 |
| 4. | AI モデルの開発・学習..... | 24 |
| 4.1 | AI モデルの全体像..... | 24 |
| 4.2 | 取得したデータ..... | 26 |
| 4.3 | AI モデルの学習プロセス..... | 29 |
| 4.4 | AI モデルの開発・学習にあたっての留意事項等..... | 30 |
| 5. | 実証の手法..... | 33 |
| 5.1 | ユースケース①:エッジ AI による通信量の低減..... | 33 |

| | | |
|-------|---|----|
| 5.2 | ユースケース②:山間部や海中等の携帯電話網不感地域における通信の確立 | 34 |
| 5.3 | ユースケース③:多種多様なデバイスに対する通信リソースの最適化 | 35 |
| 5.4 | ユースケース④:分散ネットワークにおける新たな AI ソリューションの創出 | 36 |
| 5.4.1 | 実証内容の詳細 | 36 |
| 5.4.2 | 利用技術・データ | 36 |
| 5.4.3 | 必要性・新規性 | 38 |
| 5.4.4 | 検証条件 | 40 |
| 5.4.5 | 開発・評価項目 | 41 |
| 5.4.6 | KPI/KGI | 47 |
| 5.5 | 拡張性に関する検証 | 50 |
| 5.5.1 | 実証内容の詳細 | 50 |
| 5.5.2 | 前提条件 | 51 |
| 5.5.3 | 評価項目 | 51 |
| 5.5.4 | KPI/KGI | 53 |
| 5.6 | 効果検証 | 54 |
| 5.6.1 | 実証内容の詳細 | 54 |
| 5.6.2 | 前提条件 | 54 |
| 5.6.3 | 評価項目 | 55 |
| 5.6.4 | KPI/KGI | 57 |
| 6. | 実証の結果・考察 | 59 |
| 6.1 | 全体スケジュール | 59 |
| 6.2 | ユースケース①:エッジ AI による通信量の低減 | 60 |
| 6.3 | ユースケース③:多種多様なデバイスに対する通信リソースの最適化 | 61 |
| 6.4 | ユースケース④:分散ネットワークにおける新たな AI ソリューションの創出 | 62 |
| 6.4.1 | 実証スケジュール | 62 |
| 6.4.2 | 開発・評価項目の結果 | 62 |
| 6.4.3 | KPI/KGI との比較結果 | 69 |
| 6.4.4 | 考察 | 71 |
| 6.5 | 拡張性に関する検証 | 73 |
| 6.5.1 | 実証スケジュール | 73 |
| 6.5.2 | 評価項目の結果 | 74 |
| 6.5.3 | KPI/KGI との比較結果 | 79 |
| 6.5.4 | 考察 | 80 |
| 6.6 | 効果検証 | 81 |
| 6.6.1 | 実証スケジュール | 81 |
| 6.6.2 | 評価項目の結果 | 82 |
| 6.6.3 | KPI/KGI との比較結果 | 85 |
| 6.6.4 | 考察 | 86 |

| | | |
|-----|---|----|
| 7. | 本実証の総括 | 87 |
| 7.1 | 本実証の成果・課題 | 87 |
| 8. | 参考資料 | 89 |
| 8.1 | 付録 1 F-RAG WebAPI 仕様 詳細 | 89 |
| 8.2 | 付録 2 観光客向けアプリケーション動作イメージ | 90 |
| 8.3 | 付録 3 機械学習モデルの詳細構成(TCN + Node-wise Self-Attention) | 93 |
| 8.4 | 付録 4 入力特徴量の詳細(入力次元) | 95 |

0. エグゼクティブサマリ

0.1 実証概要

本実証は、地域における「持続可能なデジタルインフラ」の構築を目的とし、AI を活用した効率的なデータ収集・活用の仕組みと、それを元にしたサービス創出を可能とするデータ基盤の開発を目指すものである。

カメラ映像データ及びそのAIによる解析結果は、人の目視作業を代替し、人物や物体の動作、位置、数、移動動線といった多様な状態を遠隔から把握できるため、監視、安全管理、施設運営のモニタリング等、幅広い分野への応用が期待される。これらが社会実装されれば、時間や場所の制約を超えた状況把握が可能となり、人手をより高付加価値な業務へ振り向けることで、社会課題の解決と新たな価値創出の両立が可能となる。一方で、映像解析を前提とした AI の開発・運用には大規模データや高速通信、計算資源が必要であり、またプライバシー・セキュリティ対策や導入・運用コストの負担も大きいことから、地域社会での実装は容易ではない。

そこで本実証では、エッジAI及び連合学習を活用の上、データを中央に集約することなく、ネットワーク負荷の削減と、高精度なAI処理を両立する手法の有効性を検証する。具体的には、北海道函館市内の主要観光スポットにおいて、エッジAIによりリアルタイムな人流データを取得し、連合学習によって地域特性を反映した混雑予測モデルを生成・更新する。更に、得られた混雑予測結果及び都市データを、クエリ先選択型 F-RAG を介して必要な範囲に限定して参照し、大規模言語モデルで回答を生成することで、対話型の情報提供を検証する。また、観光・交通分野における意思決定支援や行動誘導への有効性を検証すると共に、本基盤が、他分野・他地域へ展開可能なデジタルインフラとなり得るかを評価する。

0.2 KPI/KGI の内容と達成状況

| 番号 | 目標値 | 結果 | 達成状況 |
|------|-------------------------------|---------------|--|
| 定量評価 | 1 通信コスト 20%以上削減 | 99%削減 | ➤ 連合学習で数値データ及びモデル更新情報のみに送信対象を限定した条件では効果が大きかった。 |
| | 2 情報提供のリアルタイム性 300ms 以下 | 98%実行 | ➤ エッジ側で混雑推定を完結させ、API 取得対象を必要な地点・時間帯に絞った条件下では、遅延が安定的に抑えられ、効果が顕著であった。 |
| | 3 情報公開による混雑緩和 混雑率 20%以上の低下 | 50%低下 | ➤ 混雑実測値に加えて近未来の混雑予測や代替時間・地点の提案を併せて提示した条件下では、緩和効果が大きく現れた。 |
| | 4 混雑予測の精度 20%以上の向上 | RMSE 80%削減 | ➤ 連合学習により各地点の特性を反映した学習を行い、時間帯や曜日、天候といった外部要因を入力に含めた条件下では、改善効果が顕著であった。 |
| | 5 モデル更新時間 20%以上の短縮 | 80%短縮 | ➤ エッジ側での事前学習と連合学習によるモデル更新を組み合わせ、更新頻度を日次・バッチ処理とした条件下では、更新処理効率化による短縮効果が大きく現れた。 |

| | | | | |
|------|---|-----------------------|------------|--|
| | 6 | データ転送量の削減 学習目的を 0% | 0%達成 | ▶ ローカル処理とデータ圧縮をおこなうことで、学習を目的とした転送は 0%にできた。 |
| | 7 | AI エージェントの組み込み | 正常動作確認 | ▶ 巡回制御により取得対象を限定し、エッジ推論や分散 RAG と連携した条件下では、効果が大きく現れた。 |
| 定性評価 | 1 | インフラ基盤の需要や活用可能性の明確化 | ヒアリング取りまとめ | ▶ 都市の業務課題解決に繋がるユースケースでは、導入を前提とした実務的な関心が高く示された。 |

0.3 考察

本実証を通じて、エッジ AI、連合学習、F-RAG を中核とした AI インフラ基盤の動作が有効に機能し得ることが確認された。特に、映像データをクラウドへ集約せず、エッジ側で解析・学習を行う構成により、通信負荷の大幅な低減とプライバシー配慮を両立しながら、高精度な混雑予測が可能である点は、従来のクラウド集中型手法に対する優位性として評価できる。

技術面では、モバイル通信環境下においても安定したデータ取得・分析が可能であり、地域インフラの制約を前提とした実装の現実性が示された。また、連合学習により地域特性を反映したモデル更新が可能であることから、観光地や交通拠点毎に異なる人流特性に適應できることが確認された。

一方、運用面では、エッジ機器管理やモデル更新ルールの整理が今後の課題として挙げられるが、AI エージェントを活用した選択的データ取得は、運用負荷の抑制と拡張性確保に有効である。更に、F-RAG と大規模言語モデルを用いた情報提供により、リアルタイム性と文脈理解を踏まえた意思決定支援型の観光案内が可能となり、混雑分散や回遊促進への貢献が期待できる。総合的に、本実証は観光分野にとどまらず、防災や都市計画等への展開可能性を有する AI インフラ基盤の有効性と、社会実装に向けた運用等の課題を明確化した点に意義がある。

0.4 成果

AI と通信を組み合わせた分散型アーキテクチャが、地域における社会実装において継続利用できる点で有効であることが確認された。特に、エッジ AI を活用して映像データをローカルで解析し、数値化された結果のみを通信することで、通信負荷を大幅に抑制しつつ、リアルタイム性とプライバシー配慮を両立できる点は、地域実装における重要な知見である。

また、連合学習を用いることで、各エリアのデータを中央に集約することなく、地域特性を反映した高精度な混雑予測モデルを構築できることが確認された。これは、データ主権やセキュリティを重視する行政分野や公共領域において、AI 活用を進める上で有効な手法であるといえる。

観光分野のユースケースにおいては、リアルタイムな混雑把握や将来予測を活用した情報提供が、混雑分散や行動誘導に資する可能性が示された。更に、F-RAG と大規模言語モデルを組み合わせた観光案内では、静的情報にとどまらない、状況に応じた意思決定支援型サービスの有効性も確認できた。

加えて、モバイル通信を前提とした低帯域設計や、イベント駆動型通信、AI エージェントによる選択的データ活用は、通信インフラに制約のある地域でも耐えうる構成であり、他地域・他分野への横展開可能性を高める結果となった。

以上により、AI×通信による持続可能なデジタルインフラの実現に向けた実践的な知見を示すことができた。

0.5 課題

本分散型 AI インフラが、通信制約やプライバシー制約のある実環境においても有効に機能することが確認された。一方で、実運用及び他地域・他分野への横展開を見据えたいいくつかの課題も明らかとなった。

第一に、運用体制及びガバナンス設計に関する課題である。エッジ側で推論を完結させ、連合学習でモデル更新を行う構成は通信負荷やプライバシー面で優位性を持つ一方、エッジ機器の監視・保守、モデル更新の頻度や責任分界、障害発生時の対応ルール等、従来のクラウド集中型とは異なる運用設計が必要となる。特に、平時運用と災害時等の非常時運用をどのように切り替えるかは整理すべき重要な論点である。

第二に、費用対効果の可視化に関する課題である。本実証では通信量削減やモバイル回線前提での運用可能性といった技術的効果を定量的に確認できたが、導入判断に向けては、初期導入費用・運用費用と、人手調査の代替や業務省力化、混雑緩和といった便益を結び付けて示す必要がある。投資対効果を分かりやすく整理し、予算化に繋げることが求められる。

第三に、社会受容性と合意形成の継続的確保である。映像を外部に送信しない設計や連合学習の採用は理解促進に寄与した一方、広域展開や長期運用を前提とする場合には、取得データの内容や利用目的を住民・関係者に対して継続的に説明し、透明性を確保することが不可欠である。

以上より、本インフラの社会実装には、技術的有効性に加え、運用・費用・社会的側面を含めた総合的な設計と段階的な導入が重要となる。本実証で得られた知見を踏まえ、函館市の関係者との協議を継続しながら、実運用に向けた体制構築と制度面の検討を進めていきたい。

1. 実証の目的及び内容等

1.1 実証の背景・取り組むべき課題

本実証のテーマは、地域における「持続可能なデジタルインフラ」の構築を目指し、AI を活用した効率的なデータ収集・活用の仕組みの構築と、それを利活用したサービス創出を可能とするデータ基盤の開発である。従来の AI 開発は、大規模なデータセットを構築し、解析性能の向上に主眼を置いたものが中心であった。一方、本実証では、社会実装を見据え、費用対効果、セキュリティ対応、ネットワーク負荷といった運用面の要件にも十分に配慮した AI インフラの構築を目指すと共に、その基盤を活用した多様なサービス展開が可能なプラットフォームの実現を目的としている。

AI をはじめとするデジタル技術が社会課題解決と新たな価値創出の双方に寄与する顕著な例として、カメラ映像やセンサーデータを取得し、AI で解析した結果を活用する取組が挙げられる。これらの技術は、人による目視確認の自動化にとどまらず、目視では把握しきれない事象の検出を可能とし、物体や人の動作、位置、数、移動動線等の多様な状態を遠隔から把握できる。そのため、監視業務、安全管理、施設運営のモニタリング等、幅広い用途への活用が期待される。AI 技術の進歩と学習データの蓄積により、これらの解析精度は今後更に高度化することが見込まれ、地域社会や産業分野に浸透すれば、時間や場所の制約を受けることなく高付加価値な業務への注力が可能となり、産業の魅力向上や活性化、更には収益性の向上にも寄与する。

一方で、こうした技術を実際に活用する際には、AI 解析を実行するためのコンピューティングリソースの確保に加え、導入・保守にかかるコスト、プライバシー及びセキュリティへの対応といった技術面・費用面の負担が大きな課題となる。大量のデータをリアルタイムに処理するためには高速かつ安定した通信環境が必要であり、カメラ等のセンサーデバイスの設置条件や台数、利用する AI アルゴリズムによって求められる要件も大きく異なる。更に、対象エリアの拡大や用途の追加に伴う拡張性の確保も容易ではない。加えて、リアルタイムデータを用いた高度な AI シミュレーション等を継続的に実行する場合、通信インフラへの負荷は一層増大し、クラウド、エッジ、デバイスを含めたネットワーク全体の効率的な統合と、AI 技術との最適な組み合わせが不可欠となる。このような背景から、AI 技術の導入・活用は地域社会において依然として進みにくい状況にある。

そこで本実証では、カメラ等によるデータ取得及びその利活用を促進するため、「AI を活用したデータ収集・統合」と「ネットワーク負荷の削減」を両立する手法と、それを活用したソリューションの有効性を検証する。また、社会実装を見据えた拡張性や、他サービス・他分野への活用可能性についても併せて検証を行う。その実現に向け、上記の技術的課題を解決する複数の手法を組み合わせ、プラットフォームとして構築することで、そこで得られるデータを多様な事業者へ提供し、幅広いサービス創出に繋げることを目指す。

本実証は、函館市と連携して実施する。同市は観光業が盛んな都市である一方、近年はオーバーツーリズムの課題に加え、少子高齢化という構造的な課題も抱えている。インバウンドを含む観光客の増加や大型客船の寄港により、公共交通機関の運転手や観光地での現地スタッフ不足が顕在化しており、観光需要のリアルタイムな把握や、人員・車両配置計画の高度化に対するニーズは極めて高い。これらの課題は他の地域都市にも共通するものであり、函館市は本実証のフィールドとして高い適性を有している。

更に、他自治体への適用を見据え、本取組を函館市内の一部エリアや特定用途に限定するのではなく、対象エリア拡大時の対応や、都市計画、防災・災害時の避難計画等への応用可能性についても検討する。これにより、個別事業者のアプリケーション開発にとどまらず、地域自治体全体に展開可能なデジタル基盤としての実装を目指す。

1.2 実証の目的

本実証の目的は、低コストで持続可能な AI インフラを構築し、その基盤を活用することで多様なサービス創出を実現することである。特に、大量データの収集・処理に伴う計算資源の消費やネットワーク負荷といった課題に対し、AI を活用した効率的なアプローチ手法を検証すると共に、それらを組み合わせた分散型の基盤構成を構築する。

本実証では、データの取得から解析、蓄積、提供までを一体的に捉え、通信負荷や運用コストを抑制しながら継続利用が可能な AI インフラの実現性を評価する。併せて、当該基盤を通じて得られたデータを、観光・交通・防災等の多様な分野に携わる事業者へ提供することで、分野横断的なサービス創出の可能性を検証することを目的とする。主な取り組み項目は、以下の 4 点とする。

①データ取得、転送の最適化

以下の各ケースにおける LTE 等のモバイル通信、Wi-Fi、有線等の最適ネットワーク(コスト、転送速度等)を比較検証の上、その最適な手法を検討する。

- ・カメラ等で取得したデータをエッジボックスにて処理・クラウドへ結果のみ送信する方法
- ・映像データをクラウドへ転送の上、解析(高性能エッジボックスの設置が不可な場合を想定)する方法

②AI エージェントによるデータ収集、提供の効率化

常にデータを送信する仕様とする場合、デバイスの増加に伴いネットワーク負荷も増大すると共に、転送される側のクラウドサービスの処理リソースも増大することになる。今回、自律的な AI エージェントがデータ取得のデータベースを巡回し、AI 処理やソリューションに必要なデータのみを必要なタイミングで取得する仕組みを提供することで、無駄なデータ取得を削減し、ネットワーク負荷を削減する。またバーチャルセンサー技術によりエリア全体を推論し、設置デバイス数の最適化を行う。

③連合学習による予測モデルの AI システム構築

エッジボックスもしくは、エッジボックス同士で直接データ連携し、その中でリアルタイム学習を行い、エリア特有の混雑パターン等を反映することで、季節変動やイベント等の混雑の変動に適応し、かつ学習したモデルパラメータのみをクラウドに送信することで、ネットワーク負荷を削減する。また、クエリ先選択型 F-RAG(連邦型 RAG)」という新たな技術を導入することで、情報の分散保持とプライバシー保護に対応する。

④データ連携基盤の有効性と他展開のための拡張性検証

上記の技術を組み合わせつつ、データ連携基盤を構築する。以下の通り、観光の事例における効果を検証しながら、他都市や他分野への展開のための拡張性の検証を実施する。

1. 観光における有効性検証

観光のシーンにおける人流データの利用ケースにあてはめながら、観光客の利便性向上・周遊促進を目指した効果検証を実施する。

- ・人流データや観光バス、市電等の交通手段の人流情報を集約の上、スポットの近い未来の混雑情報を公開し、効率のよい周遊を提案する。
- ・混雑予測を利用して市電の臨時便を出した時の効果をシミュレーションの上、実施判断への有効性を検証する。

2. 他展開のための拡張性検証

今後の拡張性のために、カメラ等センサーデバイスの設置条件、台数、使用する AI のアルゴリズム等に適した要件の構成を踏まえて、データを取得する対象エリアを広げる対応について効率的な手法を検証する。更に、他分野の適用として、都市計画策定への活用や災害時の避難計画への策定にとって有効であるか確認する。

1.3 最終目標・構想イメージ

今回の実証を通じて、広範囲なエリアにおいても効率的なデータ収集、提供及び学習が可能な AI インフラ基盤を構築することの有効性を確認する。併せて、本基盤の活用により、地域市民並びに観光、交通、防災等の分野に従事する幅広い関係者にとっての利便性向上及び業務の生産性向上に寄与することを目指す。

①データ取得、転送の最適化

最適な手法の検討を経て、今後、市全体を広範囲にカバーするために、大小異なる広さのある観光スポットの追加や、路面電車の電停の追加、市バスのバス停等の追加、更には飲食店等のデータ取得に繋げて、AI に学習させることで、函館市全体の観光エリアまで拡張することを目指す。また観光だけではなく、市民の移動の最適化、災害時の影響把握等、観光と密接に絡むところへも応用できるインフラを目指す。

②AI エージェントによるデータ収集、提供の効率化

AI エージェントを活用し、市内の観光・交通分野におけるリアルタイムなデータ収集・統合・提供を自律的に最適化するシステムを構築し、データ取得箇所が増大していく際のネットワーク負荷を削減しながら、システム全体のコンピューティングリソースの最適化を図る。また、エリア全体の統合解析によりアプリやサービスシステム向けのデータ連携タイミングも制御することで、不要な通信を削減し、データ連携するサービスが増加した際にも効率的なデータ連携機能を提供する。また、バーチャルセンサーにより、エリア全体推論 AI を活用し、カメラやエッジボックス等のエッジデバイスの効率的な設置を目指す。

③連合学習による予測モデルの AI システム構築

連合学習を活用し、各所で取得したデータを集約せずに分散で処理することにより、通信負荷を低減しながら高精度な予測を実現するモデルを構築する。中央サーバで学習結果の結合と再学習を実施することでモデルの高度化を図り、市内の観光・交通における混雑予測 AI システムを構築し、高精度な混雑予測と最適な行動提案を実現する。

④データ連携基盤の有効性と他展開のための拡張性検証

1. 観光における有効性検証

観光客が混雑時間を避けて訪問するような行動変容を促すことで、混雑を平準化することや、他の観光スポット等への訪問を提案することにより回遊性の向上を図ることで、観光客の満足度を向上させ、その結果、観光消費額の向上に繋げる。また交通機関における一例として、路面電車の市電は混雑時には、臨時便を増発するが、その実施判断は、過去の実績等や経験によるもので、確証が得られにくい。リアルタイムデータによる混雑予測を提供し、その判断に適用させることで業務効率化への貢献を目指す。

2. 他展開のための拡張性検証

デバイスを設置するパターンとその手法のテンプレート化により、屋外の商業エリア等のある程度エリアカバーが必要な場所での設置追加対応や、新規でのスポット追加の際の、セットアップの手順を確立し、函館市以外の自治体への横展開が可能な形を目指す。

更に、データ連携基盤の活用として、地元のベンチャー企業や学術機関を含む外部へのデータ提供により、自社の単独な個別アプリケーション開発の基盤に留めず、広く地域の課題解決や地元の新たなサービス創出の基盤となることを目指す。

1.4 本実証の位置づけ・目標

代表的な観光スポットにおいて、データ取得及び活用を効率的に行う仕組みを構築すると共に、当該エリアに滞在する観光客や事業者にとって、取得されたリアルタイムデータや将来予測結果が実生活や業務において有効に活用できることを目指す。本実証に向けて取り組む具体的な実施内容及び目標は、以下の通りである。

①データ取得、転送の最適化

データ取得方法について、各エリアでの人流捕捉において、ターゲットとする観光スポットに応じた以下のような考え方にて実施する。

- ・人の出入りが限定される場所(五稜郭タワー等)
最低 2 箇所において、センサーデバイスを計 10 箇所以上設置する
- ・ある程度の広いエリアをカバーする必要がある場所(函館朝市の商業エリア等)
最低 1 箇所において、センサーデバイスを計 10 箇所以上設置する
- ・路面電車等の交通移動手段
最低 3 箇所センサーデバイス計 6 台以上を設置

これらの場所でのパターンの比較検証により、ネットワークコストや転送効率等を勘案した上で、最適なネットワーク構成を検討し実現することを目指す。

②AI エージェントによるデータ収集、提供の効率化

ネットワーク負荷の削減を実現するため、AI エージェントにより 70%以上のセンサーデバイスを巡回対象として、常時データを送信し続ける場合と比較しデータ転送量の 20%以上の削減を目指す。観光客の回遊性向上や混雑平準化を実現するため、リアルタイムに取得した情報を元に混雑予測、観光の最適ルート提案のモデルを確立し、アプリやサービスシステムへ即座にデータ提供ができる仕組みの構

築を目指す。バーチャルセンサーによるエリア推論 AI によるデバイス数の最適化については、低コストで持続的な情報提供基盤の確立に向け、従来の方法に対して、20%以上の個数削減を目指す。

③連合学習による予測モデルの AI システム構築

リアルタイム性があり精度の高い予測を実現するため、エッジでの分散処理モデルを構築することで、ネットワーク利用効率を向上しながら、短期(1-2 時間後)及び中期(3-4 時間後)の予測精度を実測データと比較して差分 20%以内を目指す。クエリ先選択型 F-RAG の導入により、従来の連邦型 RAG と比較して、通信量の 50%以上の削減を目指す。データ転送量の削減率は、学習を目的としたデータ転送を 0%にすることを目的に、エッジ AI によるローカル処理とデータ圧縮を実施する。

④データ連携基盤の有効性と他展開のための拡張性検証

1. 観光における有効性検証

観光客の回遊性向上や混雑平準化の実現のため、リアルタイムデータを元にした混雑情報の提供、観光ルート案内を行うアプリ等を提供することで、観光客に行動変容を促すことを目指す。また路面電車の混雑時の臨時便の増発等の判断について、その実施の判断への有効性を検証するため、リアルタイムデータによる混雑予測を利用の上、シミュレーションを行う。

2. 他展開のための拡張性検証

他自治体への展開のしやすさの実現のため、センシングとデータ処理の構成テンプレートを構築の上、正常に動作することを確認する。また、他ユースケースへの適用の実現のため、函館市の都市計画や災害時の避難計画策定のために適した形で、他事業者へのデータが提供できるようなデータ連携基盤の開発、更に必要に応じたアプリもしくはダッシュボード等の開発を実施する。

1.5 社会実装に向けたロードマップ・取り組みの発展性

社会実装を進める上では、単発的な実証にとどまらず、「事業継続性」をいかに確保するかが重要であり、自治体や市民にとってのメリットを継続的に生み出し続けられるかが鍵になると考える。特に初期段階においては、カメラやセンサーデバイスの設置に対する地域の受容性が極めて重要である。本実証のように、人々が行き交う市街地の複数スポットにおいて継続的なセンシングを実施する事例は多くなく、丁寧な合意形成と実効性のある価値提示が不可欠である。

そこで本取組では、AI を活用することで効率的なデータ取得・解析が可能な基盤を整備し、取得したデータを観光分野に限定せず、地元企業や研究機関等の多様なプレイヤーへ提供することを通じて、新たな産業やサービスの創出を促すことを目指す。これにより、データ取得、利活用、サービス創出が循環するエコシステムの形成を図り、単なるインフラ整備にとどまらない持続的な価値創出に繋げる。

そのモデルケースとして、函館市において本実証を実施し、技術的有効性のみならず、運用面・社会受容性を含めた成功事例を確立することで、他都市への横展開を視野に入れる。観光都市が抱える課題は多くの地域に共通しており、本モデルを汎用化することで、幅広い都市への適用が可能になると考える。

その実現に向けては、以下のステップとして、エッジ AI、連合学習、AI エージェント等の各手法について機能の高度化を段階的に図りつつ、より広範囲なエリアにおいても効率的かつ高度なシミュレー

ションが可能な基盤へと発展させる。これにより、混雑回避、移動時間短縮、業務効率化といった、市民や事業従事者の実生活に直接的に効果のある手段として定着させ、社会実装へと結びつけていくことを目指す。

①データ取得、転送の最適化

取得データの有効性と維持管理コストを明確にし、持続可能なインフラ提供ができる可能性を示しつつ、本件終了後も継続して環境を維持することで、効果の検証、AIモデルの改善を実施する。また、函館市のカバー範囲を広げるために、大小異なる観光スポットの追加や、路面電車の市電の電停の追加、市バスの追加、更には飲食店等のデータ取得範囲の拡張も視野に検討を進める。

②AI エージェントによるデータ収集、提供の効率化

自律的なデータ取得最適化によるインフラ維持コストの低減を達成し、データの収集箇所の拡大しながら継続的に運用することで、さらなる効率性を追求する。また、データ取得・連携モデルを確立し自治体や事業者間の情報共有を促進しながら他の地域との連携も目指す。

③連合学習による予測モデルの AI システム構築

データを分散処理で学習することにより、リアルタイム性のある精度の高い予測を実現できるモデルを活用して、環境データや交通データ等を追加してさらなる利便性の高い予測を行う。エッジボックスの性能は年々進化しており、継続して取り組むことでローカル学習モデルの精度向上を実現し、様々なシーンで利用できる未来予測システムを構築して他都市、他用途への展開を目指す。利用者(観光客や交通事業者)からの多様な問い合わせに対し、チャットボットのように自然言語で何を聞かれても適切に答えられる汎用性の高い情報提供が求められるが、観光に関連する情報提供のパターンは非常に多岐にわたり、利用者の属性や状況に応じて異なる。そのため、あらゆる情報提供パターンをあらかじめ定義・実装しておくことは現実的ではないため、AIによる柔軟な応答を目指す。

④データ連携基盤の有効性と他展開のための拡張性検証

観光スポットの混雑予測をアプリ等での公開や通知により、観光客の行動変容を促す仕組みを活用しながら、サイネージ等の手段を追加して効果を高める。また、本取り組みに効果が得られた場合、他都市との連携に向けて、歴史的な関連性(例えば、戊辰戦争等)や地域的な関連性(姉妹都市等)、都市間を結びつけるような取り組みへの発展も目指す。

また得られた人流データを観光のシーンの用途だけではなく、自治体向けの各政策策定の用途に提供することに加えて、協力頂ける民間の交通事業者や商業施設との事業者データ連携もできるような仕組みを構築する。そのために事業者の秘密データや、カメラ映像等の個人情報のやり取りも可能なように連合学習のメリットを最大限活用する。また、得られたデータを活用してサービスを創出する事業者も、エコシステム形成のため、自治体だけでなく地元に関係を持つ企業と連携することで、ベンチャー企業や大学、研究機関等なるべく多くのプレイヤー集めを目指す。

2. 業務実施体制

2.1 実証機関

| | | |
|------|----------|---|
| 代表機関 | 法人名 | ソフトバンク株式会社 |
| | 代表者氏名 | 代表取締役 社長執行役員 兼 CEO 宮川潤一 |
| | 所在地 | 東京都港区海岸 1-7-1 |
| | 業務の概要 | 移動通信サービスの提供、携帯端末の販売、固定通信サービスの提供、インターネット接続サービスの提供 |
| 構成員 | 法人名 | 国立大学法人東京大学 |
| | 代表者氏名 | 総長 藤井輝夫 |
| | 所在地 | 東京都文京区本郷 7 丁目 3 番 1 号 |
| | 業務の概要 | 大学の運営や教育・研究活動 |
| | 構成員とする理由 | AI のアルゴリズム選択やその活用方法をアーキテクトと評価の視点両方で統括することが可能であり、本事業の AI 領域を俯瞰して設計することができるため。 |
| 構成員 | 法人名 | 宇宙サービスイノベーションラボ事業協同組合 |
| | 代表者氏名 | 代表理事 神武直彦 |
| | 所在地 | 東京都中央区日本橋室町 2 丁目 1 番 1 号 |
| | 業務の概要 | 衛星データや地球上の活動を示す各種データを統合解析し、シミュレーションファーストを標語に、知識や技術・サービスを一気通貫で提供する事業協同組合 |
| | 構成員とする理由 | 様々なユースケースによる AI 活用及び統合的な AI アルゴリズムの活用に長けている。国内 8 つの大学の研究室発スタートアップの連携コンソーシアム企業であり、最先端の知見も活用できる企業であるため。 |

2.2 実施体制図

本事業は、株式会社三菱総合研究所からの委託を受け、ソフトバンク株式会社を代表機関とし、国立大学法人東京大学及び宇宙サービスイノベーションラボ事業協同組合を構成員とするコンソーシアム体制により受託・実施するものである。各構成機関は、それぞれの専門性や知見を生かし、技術検証から社会実装を見据えた実証の推進までを分担して担う。更に、本事業の実証フィールドである函館市が協力機関として参画し、実施支援を行う。特に、住民向けのデジタル施策の企画及び運用を所管する企画部地域デジタル課を中心として、観光部、交通部、港湾部をはじめとする関係部署と緊密に連携しながら、本実証を進める(詳細は以下の図 2.2-1 の通り)。

データ取得やその利活用にあたっては、単なる技術実証にとどまらず、地域社会への影響や受容性を十分に考慮する必要がある。このため、地域企業や住民との合意形成については函館市が主体的に担い、事業内容や期待される効果について丁寧な説明を行うと共に、地域課題の解決に繋がる具体的な活用イメージを共有することで、理解促進を図っている。

本実証は、観光分野を起点としつつ、防災、交通、港湾等の他分野への波及も視野に入れ、地域における実サービスへと接続することを前提とした施策として位置付けている。函館市とコンソーシアム各機関が役割分担しながら協調して取り組むことで、地域特性に即した実証を推進すると共に、将来的な本格展開や他地域への横展開に資するモデルの構築を目指している。

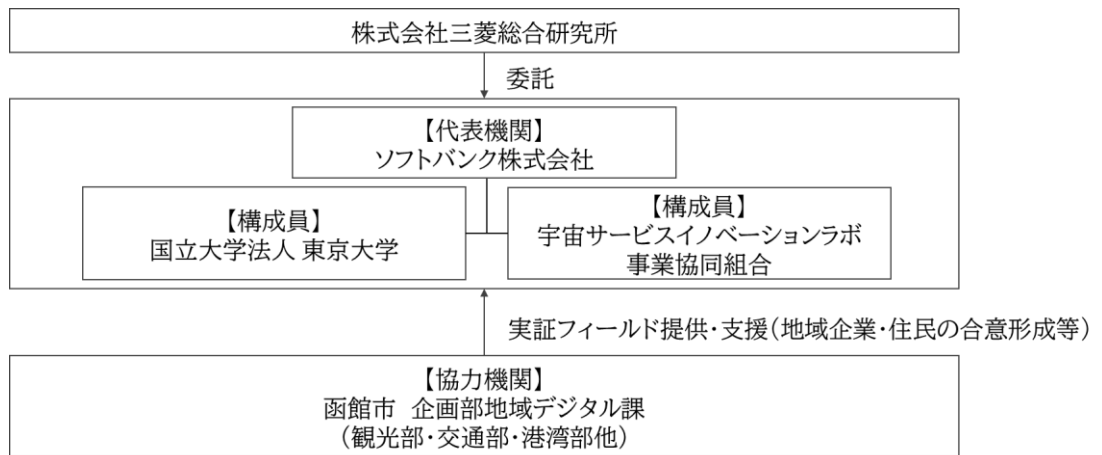


図 2.2-1 実施体制図

3. 通信システム及び AI システムの構築

3.1 通信システム及び AI システムの全体像

本実証では、フィールド側に設置したカメラと繋がるエッジボックスをクラウド環境へ接続し、エッジ AI による解析結果を保存・蓄積するデータベースを構築すると共に、API を介してアプリケーションとリアルタイムにデータを連携できるシステム全体構成を設計・構築した。(図 3.1-1)併せて、エッジ AI による映像解析、連合学習、AI エージェント、並びにクエリ先選択型 F-RAG(連邦型 RAG)といった先進的な AI 技術の有効性を検証可能とするため、システム及びネットワーク構成を設計している。

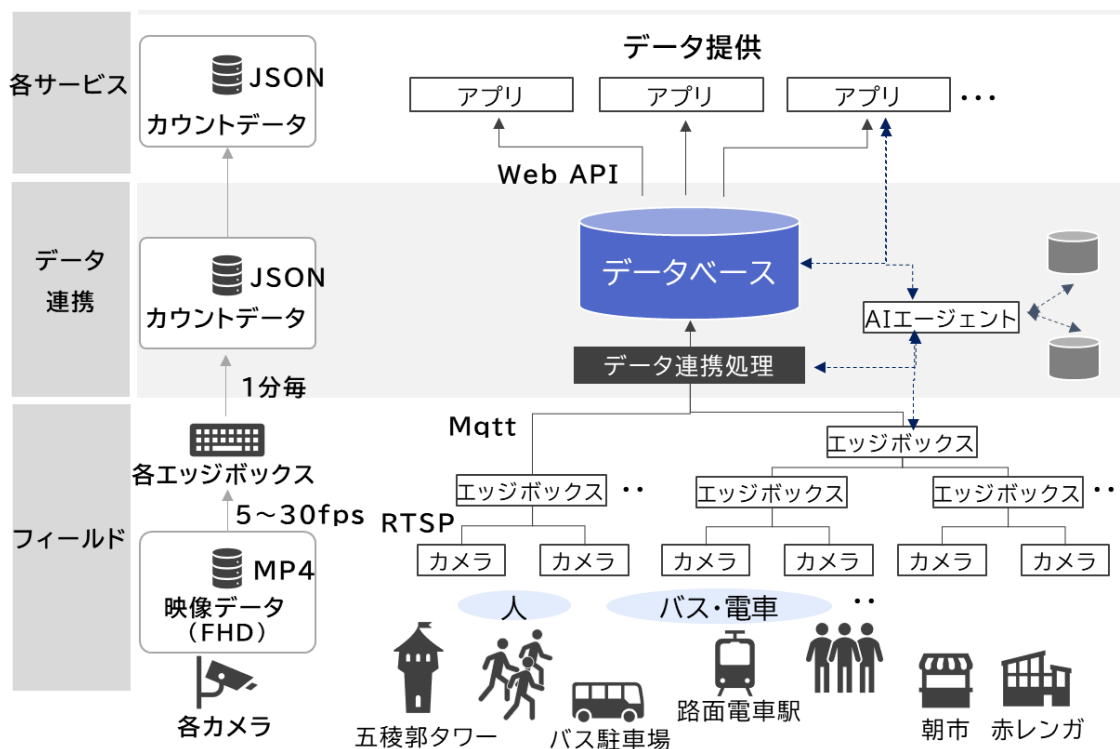


図 3.1-1 システム全体像イメージ

①フィールド側(エッジ環境)の構成

フィールド側では、カメラとエッジボックス間の通信を現地内の閉域通信環境とし、カメラ及びエッジボックスを PoE スイッチで接続することで LAN を構成している。エッジボックスには USB ドングル型の LTE モジュールを接続し、既設の有線回線に依存することなく、モバイルネットワークを介してクラウド環境へ接続する構成とした。エッジボックス上にはエッジ AI を実装し、カメラ映像から人数カウントや人流解析等の推論処理をローカルで完結させる。クラウド環境へ伝送するのは、推論結果として得られた解析済みのカウントデータのみとし、映像データや画像フレームをクラウドへ転送する処理は行わない。これにより、通信負荷を最小限に抑えると共に、個人情報保護及びプライバシー配慮を前提とした設計を実現している。

②クラウド環境の構成

クラウドネットワーク上には、推論結果データの蓄積及び大規模な解析処理を行うためのストレージ及びサーバを配置し、エッジネットワークとは VPN を介してセキュアに接続している。クラウド上に蓄積されたデータは、API を通じてアプリケーションや外部サービスと連携され、リアルタイムな可視化や分析に活用される。

③AI 技術の構成と検証内容

AI 技術としては、上述のエッジ AI に加え、連合学習、AI エージェント、クエリ先選択型 F-RAG の活用について検証を行った。混雑予測においては、各観光スポットに設置されたエッジボックスがローカルで人流データを学習し、各エリアで学習した AI モデルのパラメータのみを統合する連合学習方式を採用している。これにより、各エリアの季節性、時間帯、イベント等の特性を反映した個別最適化モデルを構築しつつ、中央クラウド側で全体として高精度な未来予測を実現することを目指している。

また、市街地に分散配置された各エッジボックスから取得される人数カウントや混雑度等の人流解析結果や、クラウドに蓄積された過去人流データや予測結果については、自律型 AI エージェントを活用してリアルタイムに収集・統合する仕組みを検証した。混雑状況や人流データを効率的に取得するため、巡回ルート最適化アルゴリズムを適用し、混雑エリアを回避しながら必要なデータを収集することで、通信量及び処理負荷の最適化を図っている。

更に、本実証で得られるリアルタイムな混雑予測結果は、それ自体を可視化するだけでなく、観光案内や行動提案といった、より高度なアプリケーションへ発展させる価値を有する情報である。このため、本実証では、混雑予測結果を活用した観光案内エージェントを開発した。観光案内エージェントでは、クエリ先選択型 F-RAG 技術を用い、分散して管理されている混雑予測データや都市データの中から、利用者の質問や状況に応じて必要なデータのみを選択的に参照する構成としている。これにより、データを一元的に集約することなく、効率的かつ安全に最新情報を活用できる。また、本エージェントでは、大規模言語モデルとして GPT を利用し、同モデルが有する観光地の一般情報や移動・行動に関する知識と、実証で得られたリアルタイムな混雑予測結果を組み合わせることで、単なる情報提示にとどまらない、状況に応じた観光案内や意思決定支援を行うことを目指している。

以上の構成により、本実証では、通信負荷の低減、プライバシー配慮、分散環境下での高精度 AI 処理を両立する AI インフラ基盤の有効性を検証可能なシステム全体構成を実現した。

3.2 通信システム及び AI システムの設置場所

データ取得のための機器設置場所については、函館市内の主要な観光スポットを対象に、函館市観光課と連携しながら現地調査を実施し、訪問者数の多さや時間帯毎の混雑状況、観光客の移動動線等の観点から総合的に検討を行った。単に人が集まる地点を選定するのではなく、観光客の行動特性や回遊の流れを把握することで、実証において有効な人流データを取得できる箇所を選定することを重視した。また、設置スポットの選定にあたっては、本実証に限らず、今後の他地域・他分野への横展開や汎用的な設計指針として活用できるよう、設置場所の特性に応じて以下の 3 つのパターンに区分し、それぞれの考え方を整理した。

A) 人の出入りが限定される場所(入退出口やチケット売り場等を有する観光施設)

入場口やチケット売り場といった人の流入・流出が集約される箇所を中心に設置を行った。これにより、入場者数や滞留状況を比較的高い精度で把握することが可能となり、施設単位での混雑状況やピーク時間帯の把握に適したデータ取得が可能となる。

B) ある程度の広いエリアをカバーする必要がある場所(屋外の市街地観光スポット等)

屋外の市街地エリアでは、単一地点での計測では人流の全体像を把握することが困難であるため、エリア内の主要な通りや交差点等の複数箇所に分散して設置する構成とした。これにより、エリア内での人の流れや滞留傾向、時間帯による変化を面的に捉えることが可能となり、回遊促進や混雑緩和施策の検討に資するデータ取得を目指した。

C) 路面電車等の交通移動手段に関わる場所(各観光地の最寄り駅や停留所等)

交通結節点となるスポットでは、路面電車の乗降口や乗車待ち列が形成されやすい箇所を中心に設置を行った。これにより、観光客の移動タイミングや集中時間帯、交通手段利用時の混雑状況を把握することが可能となり、観光地と交通分野を横断した混雑対策や運行計画検討への活用を想定している。

以上を踏まえ、実際に各場所に必要な機器を設置した(詳細は図 3.2-1 の通り)。このように設置場所を特性別に整理し検討することで、実証におけるデータ取得効果を最大化すると共に、今後、他都市や異なる分野に展開する際にも適用可能な設置指針の整理を行うことができた。



図 3.2-1 函館市内の主要観光地における設置場所

実際の機器設置にあたっては、各スポットにおける人流の規模や時間帯毎の変動、観光客の移動動線を踏まえ、現地調査結果を基に設置位置及び台数を検討した。その結果、以下の表 3.2-1 に示す通り、各スポット毎に必要な機器を設置し、観光地における混雑状況を把握・可視化するためのデータ取得を実施した。これにより、単一地点での計測にとどまらず、観光地全体の混雑傾向や人の流れを把握することを目指した。

設置する機器については、本実証に限らず、将来的な本格導入や他地域への展開を見据え、比較的安価であり、かつ現時点で安定的に入手可能なモデルを基本とする方針とした。この観点から、カメラには i-PRO (iPro) 社製のネットワークカメラを、エッジボックスには NVIDIA 社製 Jetson Orin Nano を基本構成として選定した。これらの機器は、屋外設置に耐えうる信頼性を有しつつ、エッジ AI による映像解析を実施するうえで十分な性能を備えており、コストと性能のバランスに優れている点を評価している。尚、本エッジボックスは、高速な AI 推論を実現するために必要な GPU を搭載しており、小型かつ低消費電力であるため、設置場所の制約が大きい市街地や屋内外混在環境においても導入しやすい点が特長である。加えて、他の GPU 搭載機器と比較して価格が比較的安価であり、国内在庫が安定して確保できる点も、実証事業及び将来的な横展開を見据えた際の優位性として評価した。また、将来的な選択肢拡大や構成最適化の検討を目的として、Luxonis 社製のカメラ及び Apple 社製 Mac mini についても、比較検証用途として併せて活用した。これにより、異なるハードウェア構成における処理性能、運用性、消費電力、設置性等の違いを把握し、用途や設置条件に応じた最適な機器選定に向けた知見を蓄積することができた。

以上のように、実証においては、現実的なコスト制約や調達性を考慮した機器構成を採用すると共に、複数の選択肢を比較検証することで、観光地の混雑情報活用に資する実用的かつ汎用性の高いデータ取得環境を構築した。

表 3.2-1 設置スポット毎の設置物と捕捉対象

| パターン | 設置スポット | カメラ | エッジボックス | 捕捉対象(取得データ) |
|------|-----------------------|------|---------|----------------------------|
| A | 五稜郭タワー | 14 台 | 8 台 | 入退人数、チケット・展望フロア待機列数、観光バス台数 |
| A | クルーズターミナル | 3 台 | 2 台 | ターミナル流入出数、朝市方面流入出数、駐車場バス台数 |
| B | 函館朝市エリア | 11 台 | 6 台 | 滞留人数、エリアへの流入出数 |
| C | 電停駅(五稜郭公園前・函館駅前・十字街駅) | 6 台 | 3 台 | 乗車・降車人数(方向別) |

設置機器の構成については、エッジボックスにおける AI 推論処理及びデータ処理の負荷を考慮し、エッジボックス 1 台あたり接続するカメラは最大 2 台までを基本構成とした。この構成により、リアルタイム性を確保しつつ、安定した処理性能を維持すると共に、将来的な増設や配置変更にも柔軟に対応できる設計としている。カメラとエッジボックス間の通信については、建物内に設置するケースを想定し、屋内で LAN ケーブルによる配線を行った。カメラへの電源供給及び通信を一本の LAN ケーブルで実現するため、PoE(Power over Ethernet)対応スイッチを利用し、配線作業の簡素化と設置コストの低減を図った。これにより、電源工事を最小限に抑えつつ、安定した通信環境を構築している。エッジボックスから外部への通信については、設置場所において既設ネットワークの流用が困難であることや、屋外環境で新たに有線のインターネット回線を構築することが現実的でないケースが多かった。このため、エッジボックス単体で外部ネットワークへ接続可能な構成とすることを前提とし、USB 型の LTE ドングルを用いたモバイル通信方式を採用した。これにより、設置場所に依存せず柔軟に通信環境を確保でき、実証期間中の運用安定性を高めている。

データの収集・蓄積及び外部システムとの連携を行う基盤としては、クラウドサービスとして Microsoft Azure を利用した。エッジ側からは、モバイル通信網を介してインターネット経由で Azure 上のデータ連携基盤へ接続する構成としている。通信プロトコルには、軽量で IoT 用途に適した MQTT に TLS 暗号化を施した MQTTS を採用し、データ送受信時の通信経路の暗号化を行うことで、セキュリティを担保した。更に、エッジボックスへのリモートアクセスや管理操作についても、VPN を介したアクセス制御を行うことで、第三者による不正アクセスを防止している。このように、通信経路及び管理アクセスの両面からセキュリティ対策を講じることで、実証環境においても実運用を想定した安全性を確保している。

本構成(図 3.2-2)は、限られた設置条件や通信制約の中でも安定した運用が可能であり、将来的な機器増設や他エリア・他都市への展開においても再利用性の高い、汎用的なエッジ AI インフラ構成となっている。

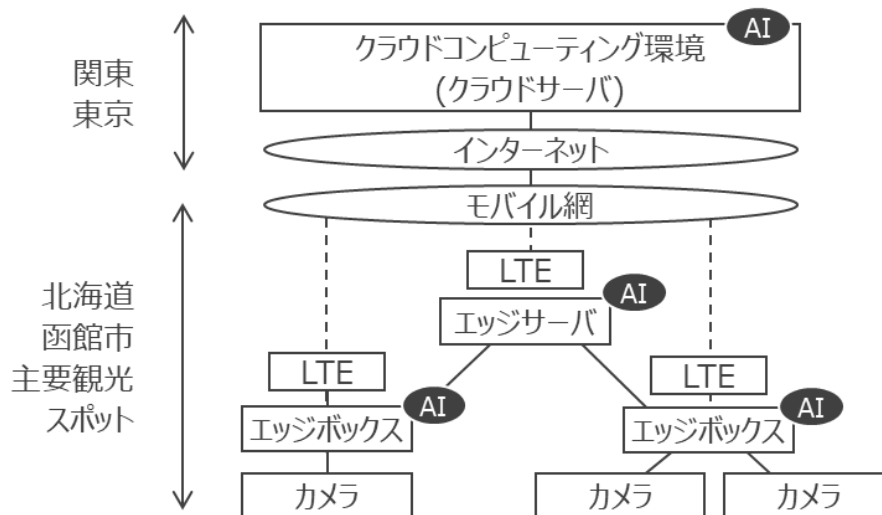


図 3.2-2 ネットワーク構成図

AI 技術については、混雑予測を行う AI システムとして、エッジデバイス、エッジサーバ、クラウドサーバからなる三層構成を採用し、混雑度データ及び AI モデルを循環させるアーキテクチャを構築した(システム構成全体は図 3.2-3 の通り)。本構成により、リアルタイム性、通信負荷低減、プライバシー保護、並びにモデル精度の継続的向上を同時に実現することを目指している。

まず、エッジデバイス層では、観光地や商業施設等に設置したカメラ映像を用いて、画像認識技術による人物検出及び人数カウントを実施する。映像そのものを外部へ送信することなく、デバイス上で処理を完結させることで、個人を特定し得る情報の流出を防止すると共に、通信量を最小限に抑えている。これらの処理結果である人数カウント情報を基に、時間帯毎の混雑度を表す時系列データを生成し、併せてカメラやデバイスの稼働状況等の状態情報を取得する。また、必要に応じて、エッジデバイス上での軽量のローカル学習により更新されたモデルパラメータを生成する。

次に、エッジサーバ層では、複数のエッジデバイスから送信された混雑度データや状態情報を集約し、天候情報、イベント開催情報、交通状況等の外部データと統合する。これにより、単一地点の情報にとどまらず、地域全体の状況を考慮した特徴量を生成することが可能となる。これらのデータを用いて、地域単位でのモデル学習及び推論を行い、短期的な混雑予測や傾向分析を実施する。ここで得られた集約済みの混雑度データ、学習済みモデルパラメータ、統合後の特徴量は、クラウドサーバ層へ送信される。

クラウドサーバ層では、複数のエッジサーバから受信したモデルパラメータを用いて、連合学習の手法により統合処理を行う。各地域の生データを直接収集・集中管理することなく、モデルパラメータのみを統合することで、プライバシーを確保しつつ、全体として最適化されたグローバルモデルを構築・評価する。併せて、予測精度の評価やモデルの安定性確認を行い、必要な調整を加える。

更新されたグローバルモデルのパラメータや学習設定情報、並びに予測精度に関するフィードバックは、再びエッジサーバ層へ配信される。エッジサーバではこれらを反映することで、地域特性に適応しつつ、全体としての性能を維持・向上させることが可能となる。このように、エッジデバイス層、エッジサーバ層、クラウドサーバ層が役割分担しながら連携することで、通信負荷を抑制しつつ、継続的に精度が向上する混雑予測 AI システムを実現する。

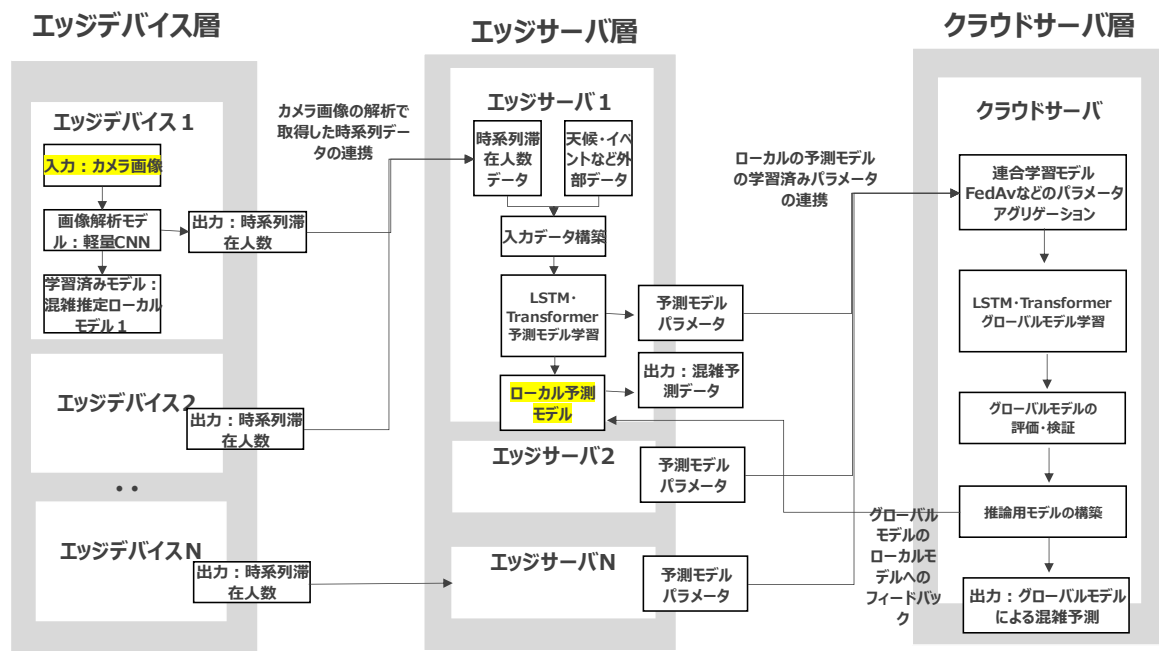


図 3.2-3 連合学習のシステム構成イメージ

映像データから得られた解析結果は、フレーム単位で処理された後、エッジボックス内に構築したローカルデータベース(SQLite)へ保存される。映像そのものは保存せず、人数カウントや通過人数等の解析結果のみを記録することで、プライバシーに配慮すると共に、エッジデバイスの記憶容量や処理負荷を抑制している。フレーム毎の解析結果データは、一定時間分を一時的に蓄積する運用とし、その保存間隔や集計周期は、カメラの設置位置や AI 処理タスクの特性に応じて柔軟に設定した。例えば、人通りが多く混雑変動の大きい場所における人数カウント処理については、変化を細かく捉えることを目的に、30 秒間隔でデータを保存する設定とした。一方、特定のラインを通過した人数を把握するクロスライン人数カウントについては、通信頻度や集計処理負荷を考慮し、1 分間隔での保存・集計とした。これらの解析結果は、1 分毎に集計処理を行った上で、データ連携処理機能を介して外部システムへ送信する構成としている。送信時の通信プロトコルには MQTT を採用した。MQTT は軽量で通信オーバーヘッドが小さく、モバイル通信等の不安定なネットワーク環境下においても安定的にデータを送受信できる点、並びに多対多の通信モデル(Publish/Subscribe)をサポートしている点において、本実証の要件に適していると判断した。

データ転送にあたっては、データ管理及び後続処理を容易にするため、データ取得場所や機器を識別するための各種 ID、データベースの主キー情報等を付与したうえで送信している。これにより、クラウド側でのデータ統合や時系列分析、トラブル発生時の追跡性を確保している。送信先としては、主にクラウド上に構築したデータベースや、リアルタイム性が求められる処理を行うアプリケーション等があり、データの内容や利用目的に応じて適切にルーティングを行う構成とした。データ連携先の仕様に応じて、Webhook 型 API や MQTT 等のプロトコルを使い分け、柔軟かつ拡張性のあるデータ連携を実現している。このように、エッジ側での解析・集計・一時保存と、必要最小限のデータを適切な周期で外部へ連携する設計とすることで、通信負荷の低減、リアルタイム性の確保、プライバシー保護を両立したデータ取得・連携基盤を構築した。

クラウド上のデータベースでは、デバイスの構成や設置場所等を定義するマスターデータと、デバイス

から送信される人数カウント等の数値情報であるトランザクションデータの 2 種類のデータを扱う。本実証では、これらのデータは内容、データサイズ、更新頻度、利用目的といった特性が大きく異なることから、用途に応じて異なるデータベースシステムを採用する構成とした。

まず、マスターデータについては、データの整合性を担保しつつ、柔軟な検索や参照が求められることから、リレーショナルデータベース(RDB)を採用した。特に、カメラの設置場所や観光地のエリア範囲等、地理的な座標情報を扱う必要があることから、データベース製品としては PostgreSQL に空間情報処理機能である PostGIS を組み込んだ構成を採用した。一方、トランザクションデータについては、将来的に多様な種類のデバイスから、異なる形式のデータが連携されることを想定している。また、性能面の特性としては、高頻度でデータが追加される一方、既存データの更新はほとんど発生しないという特徴がある。これらを踏まえ、データ形式の柔軟性と高速な書き込み性能の両立が可能な NoSQL データベースが適していると判断し、本実証では Azure Cosmos DB for NoSQL を採用した。マスターデータについては、本実証における構成を表現するための最小限の情報として、以下の 3 種類のテーブルを作成した。

- ・テナント情報テーブル:実証に関わる組織や利用主体を管理
- ・ロケーション情報テーブル:観光地やエリアの地理的範囲を管理
(地理的な範囲は PostGIS 形式のポリゴンデータとして定義)
- ・カメラ情報テーブル:カメラの設置位置及び設置方向等の情報を管理

また、トランザクションデータについては、運用性及び将来的な拡張性を考慮し、プラットフォーム側では受信したデータを可能な限り加工せず、そのまま保存する方針とした。受信したデータは、データの種類毎に Cosmos DB のコンテナを分離して保存し、データ内容は編集せずに蓄積することで、後続の分析や用途拡張に柔軟に対応できる構成としている。

API 作成にあたっては、マスターデータ及びトランザクションデータを外部システムやアプリケーションから利用可能とするため、それぞれに対応する形で構築した。マスターデータ提供用としては、ロケーション情報取得 API、カメラ情報取得 API を作成し、これらの API を用いて、ロケーションの地理的範囲やカメラの設置場所を地図データ上に可視化できることを確認した。トランザクションデータをアプリケーション上で表示・可視化する目的として、人流に関するラインクロスカウント取得 API 及びエリアカウント取得 API を作成した。これにより、リアルタイム性が求められる混雑状況の把握や、時系列での人流分析をアプリケーション上で実現している。

データ通信要件として映像解析後にクラウドへ送信されるデータは、映像そのものではなく解析結果のテキストデータとし、1 分間あたりのデータサイズは約 5KB 以下であることを前提とする。本実証では、1 台のエッジボックスに最大 3 台のカメラを接続し、それぞれの解析結果を集約して送信する構成を想定している。これらを踏まえ、通信要件は以下の通りとした。

- ・アップリンク(エッジ → クラウド)
解析結果データ送信を想定の上、バッファ処理を考慮し、5kbps 以上の通信帯域を確保する
- ・ダウンリンク(クラウド → エッジ)

AI エージェントからの制御情報配信やモデル更新等を想定の上、10Mbps 程度の通信速度を確保
この構成により、モバイル通信回線等の帯域制約がある環境下においても、安定したデータ連携が可能となる。

セキュリティ対策としては、外部からの不正アクセス防止及びノード間通信の安全性確保のため、通

信の暗号化及びネットワークの分離を基本方針とし、VPN の導入を必須とした。

・デバイス～クラウド間通信

エッジからクラウドへの通信には、MQTT に TLS 暗号化を施した MQTTS (TLS1.2) を採用し、通信経路全体を暗号化することで、外部ネットワーク区間における盗聴・改ざんを防止する。またクラウド側では、接続元となるデバイスのグローバル IP レンジを制限し、不正なデバイスからの接続を防止する。カメラ～エッジボックス間通信映像取得はストリーミング方式とするが、ID / パスワードによる認証を必須とした。LAN 区間であっても、認証情報の管理を行い、内部ネットワークにおけるセキュリティにも配慮した設計とした。

・API 連携及び連合学習基盤

連合学習基盤やアプリケーションとの API 連携においては、トークンによる認証方式を採用した。併せて、IP フィルタリングを実施し、許可された通信元以外からのアクセスを遮断する。これにより、不正アクセスやなりすましに対する多層的な防御を実現している。

3.3 通信システム及び AI システムの構築・運用に要した費用

今回の実証事業環境構築及び運用に要した費用は、総額 2,232 万円である。本費用には、観光地及び市街地エリアにおけるデータ取得環境の構築から、エッジ AI による解析、データ連携基盤の運用、並びに情報可視化のためのアプリケーション開発まで、実証に必要な一連の要素が含まれている。

まず、機器設置に関するイニシャル費用として、合計 4 スポットにおけるカメラ 34 台分の設備・機器購入費として計 522 万円を要した。これらの機器は、屋外設置に耐えうる耐候性を備えると共に、将来的な増設や他エリアへの転用も可能な汎用的仕様を前提として選定している。また、これら機器の設置工事費として計 770 万円を要した。市街地エリアでの設置であることから、安全確保や景観への配慮、既存インフラとの調整が必要となり、夜間作業や交通規制への対応、設置箇所毎の個別調整等を含めた工事内容となっている。これにより、実証期間中の安定運用と安全性の確保を図った。

次に、エッジ環境において映像データから人流解析を実施するためのエッジ AI 実装費用として 890 万円を要した。本費用には、カメラ映像をクラウドへ転送することなくローカルで処理するための AI モデルの実装、エッジデバイス上での最適化、並びにプライバシー保護を考慮した映像解析処理の設計・調整が含まれている。これにより、通信負荷の低減とリアルタイム性の両立を図っている。

更に、取得したデータを連携・管理するためのデータ連携基盤の運用に係るクラウドサービス利用料として 50 万円を計上した。本費用は、実証期間中における最小構成での利用を前提としており、将来的なスケール拡張を見据えた基盤設計の検証も兼ねている。

加えて、解析結果や人流情報等を関係者が直感的に把握できるようにするため、データ可視化・表示を目的としたアプリケーション開発費として 300 万円を要した。本アプリケーションは、実証関係者や行政担当者による利用を想定し、リアルタイム性や操作性を重視した設計としている。

| | | | |
|------------------------------|----------------------|--------|----------|
| カメラ・エッジボックス 設置のイニシャルコスト | ① 五稜郭タワー(カメラ 14 台) | 設備・機器費 | 227 万円 |
| | | 設置工事費 | 268 万円 |
| | ② 電停 3 駅(カメラ 6 台) | 設備・機器費 | 89 万円 |
| | | 設置工事費 | 136 万円 |
| | ③ クルーズターミナル(カメラ 3 台) | 設備・機器費 | 30 万円 |
| | | 設置工事費 | 66 万円 |
| | ④ 函館朝市エリア(カメラ 11 台) | 設備・機器費 | 176 万円 |
| | | 設置工事費 | 300 万円 |
| [小計] 4 スポット カメラ 34 台 | 設備・機器費 | 522 万円 | |
| | | 設置工事費 | 770 万円 |
| エッジデバイス映像解析 実装開発のイニシャルコスト | エッジ AI 解析モジュール開発 | | 890 万円 |
| データ連携基盤運用 | クラウドサービス利用料・6 か月間 | | 50 万円 |
| アプリケーション開発 | 開発初期費用 | | 300 万円 |
| 合計 | | | 2,532 万円 |

※記載の金額は実証中における参考価格であり実装時は諸条件に応じて大きく変動する可能性がある。

※設計・設置・運用、エッジ AI、連合学習・AI エージェント開発・チューニング等の人件費は含んでいない。

3.4 通信システム及び AI システムの構築にあたっての留意事項等

本実証における主な留意点として、今後の持続的な運用及び他地域への展開を見据え、機器設置やデータ取得、並びにその利活用に対する地域における受容性の向上を重要な前提条件とした。特に、市街地エリアにおいてカメラやセンサーデバイスを設置するにあたっては、技術的な実現性のみならず、地域住民や関係事業者の理解と協力を得ることが不可欠である。

このため、函館市内の関係者に対する設置許可や協力の承諾を得る過程においては、単なる観光客向けの情報提供にとどまらず、観光混雑の緩和、防災・危機管理の高度化、都市計画や交通施策への活用といった地域課題の解決にどのように寄与し得るかについて、将来的な可能性を含めて丁寧な説明を行った。これらの説明は、函館市役所職員と連携し、行政としての視点や既存の地域施策との整合性を踏まえた形で実施することで、関係者からの理解促進を図った。

また、実際の機器設置にあたっては、安全性の確保及び景観への配慮を最優先事項とし、歩行者や車両の通行に支障をきたさないこと、強風や降雪といった地域特有の気象条件にも耐えうることを前提とした設計・施工を行った。更に、将来的な増設や他エリアへの展開を見据え、特殊な工事や高額な専用機材を必要としない、比較的安価かつ汎用性の高い設置手法を採用することで、初期投資を抑えつつ、実証終了後の継続利用や横展開を容易にすることを目指した。

本実証を通じて得られた知見は、技術的な有効性に加え、地域との合意形成や社会受容性を高めるためのプロセスそのものとしても重要な成果であり、今後の AI インフラ基盤の社会実装に向けた指針として活用していく。

①関係者説明・受容性の獲得

機器設置に際しては、市街地屋外への設置候補として、まず函館市が所有する街路灯の活用を検討した。関係部局への確認・許可取得を進めた結果、街路灯の電源仕様が夜間のみ電力供給であり、日中帯の電源確保が困難であることが判明した。このため、計画を見直し、道路周辺の店舗や施設への設置へと方針変更を行った。その結果、設置場所の所有者との調整や、電源確保のための事前調査に想定以上の時間を要することとなり、機器設置スケジュールにおける遅延リスクを考慮した全体計画の策定の重要性が明らかとなった。電源については、新規敷設を行う場合、工期の長期化やコスト増に繋がるため、各管理者と調整を行い、既設電源を利用することについて承諾を得た。人流を計測したい箇所を踏まえつつ、電源が安定的に確保できる場所を優先的に選定し、機器配置を行った。

②機器選定・設置における工夫

観光地としての景観への配慮を重視し、特に歴史的景観を有する地域においては、カメラを取り付けるポールの色を周辺建物と同系色とする等、外観への影響を最小限に抑える工夫を行った。また、エッジボックスを収容するキャビネットについても、目立たない位置への配置や、周囲の景観と調和する設計を採用した。路面電車駅等では、乗車待ちを行う人の頭上にカメラを設置せざるを得ない場所については、落下防止対策を含む安全対策を十分に考慮した設置を行った。使用機器の中には防水対応が十分でないものもあるため、キャビネット選定にあたっては防水性能の基準を満たす製品を採用した。更に、冬季においては外気温が機器の動作保証温度(0℃以下)を下回る環境となることが想定されたため、キャビネット内にヒーターを設置し、安定した動作環境の確保を行った。尚、各機器選定にあたっては、今後の汎用性のために、なるべく安価でかつ国内入手しやすいモデルを検討した。

③人流データ取得精度の確保

カメラ映像解析により観光スポット単位の人流データを取得するにあたり、人物同士の重なり(被り)が検出精度低下の要因となることから、設置高さは原則として 2.5m 以上とした。また、撮影角度によって被りが発生するケースを避けるため、可能な限り斜め上方からの撮影角度を確保し、設置時には実際に人がカメラ前を通行する状況を確認しながら、最適な角度に調整・固定を行った。設置条件上、観測対象エリア内にカメラを設置できない場合については、光学ズーム機能を備えた望遠カメラを用い、離れた場所からの映像取得を行った(例:五稜郭前駅周辺)。これにより、設置制約がある環境においても、必要な人流データを取得できる構成とした。

4. AI モデルの開発・学習

4.1 AI モデルの全体像

対象となる観光エリア内の主要な観光スポット及び、それらを結ぶ公共交通機関である主要市電駅における混雑予測の精度向上を図るため、本実証では、通信負荷を最小限に抑えつつ、高精度な混雑予測を実現することの両立を設計方針としている。この方針のもと、以下の主要な AI モデル・アルゴリズムを役割分担させて活用する。

① 混雑予測モデル(時系列データ解析)

混雑予測の中核となるモデルとして、LSTM(Long Short-Term Memory)や Transformer を用いた時系列予測モデルを採用する。これらのモデルは、過去の人流データや混雑度データの時間的な変化を学習することに優れており、観光地や交通拠点における混雑状況の推移を高い精度で予測することが可能である。

本モデルでは、観光スポットや市電駅毎の過去データを学習対象とすると共に、イベント開催情報、天候条件、曜日・時間帯等の外生要因を特徴量として組み込むことで、突発的な需要変動や季節性を考慮した動的な混雑予測を実現する。これにより、単純な過去トレンドに依存しない、実運用に耐えうる予測モデルの構築を目指している。

② 連合学習モデル(分散型データ学習)

複数の観光エリアや交通拠点に分散して配置されたエッジデバイスにおいて、Federated Averaging(FedAvg)を用いた連合学習を実施する。各エッジサーバでは、ローカルで取得した人流データや混雑度データを用いてモデルのローカル学習を行い、その結果として得られた学習済みモデルパラメータのみをクラウドへ送信する。

この方式により、生データを外部へ送信することなくモデルの性能向上が可能となり、プライバシーやデータの機密性を確保しつつ、通信量の大幅な削減を実現する。また、各観光エリアや交通機関毎の利用特性をローカル学習によって反映することで、地域特性に最適化された予測を行いつつ、クラウド側では複数拠点の学習成果を統合した全体最適なグローバルモデルを構築する。(図 4.1-1)

③ エッジ AI によるデータ前処理・圧縮

映像データに関しては、カメラで取得した映像をそのままクラウドへ送信するのではなく、エッジ AI によるローカル前処理を行う。具体的には、畳み込みニューラルネットワーク(CNN)を用いて人物検出や人数カウント等の解析をリアルタイムで実施し、混雑度を示す数値データや特徴量のみを抽出する。

更に、不要なデータのフィルタリングや圧縮処理をエッジ側で行うことで、クラウドへ送信するデータ量を最小限に抑え、ネットワーク帯域の最適化を図る。この構成により、モバイル通信等の制約のあるネットワーク環境下においても、安定したデータ連携とリアルタイム性を確保することが可能となる。

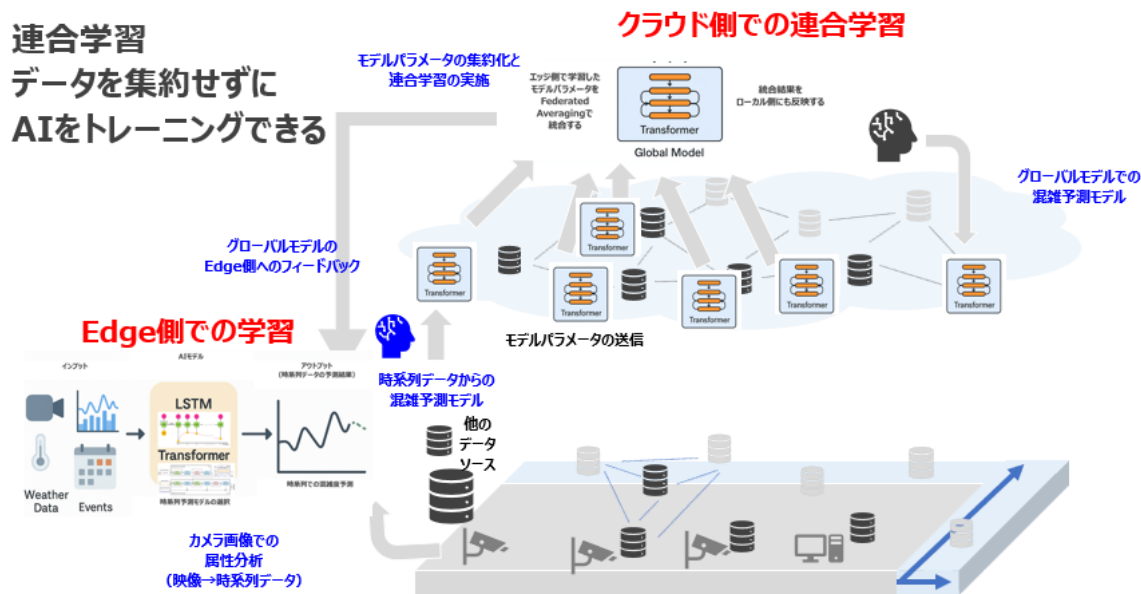


図 4.1-1 混雑予測のための連合学習全体イメージ

更に、本実証において得られた各観光エリアにおける混雑予測結果を有効活用し、観光客にとって実用的かつ分かりやすい情報提供を行うことを目的として、クエリ先選択型 F-RAG (Federated Retrieval-Augmented Generation) を用いた観光案内エージェントを構築する。本観光案内エージェントは、大規模言語モデル(LLM)が有する一般的な観光情報や知識に加え、分散して管理・保存されている混雑予測データ、交通情報、天候情報等のリアルタイム性を有するデータソースを必要に応じて参照する構成とする。クエリ先選択型 F-RAG を採用することで、ユーザーからの質問内容に応じて、参照すべきデータソースを動的に選択し、必要最小限のデータのみを取得・統合した上で回答を生成することが可能となる。

これにより、すべてのデータを一元的に集約・同期することなく、各データの管理主体や更新頻度、機密性を維持したまま、最新性の高い情報を活用した応答が実現できる。また、混雑予測データについては、観光エリア単位・時間帯単位での予測結果を参照することで、「今混んでいる場所を避けたルート案内」や「比較的空いている観光スポットの提案」といった、状況に応じた柔軟な情報提供が可能となる。観光案内エージェントは、ユーザーの質問意図を解析し、一般的な観光案内が適している場合、リアルタイムの混雑状況や予測を踏まえた回答が必要な場合、複数のデータソースを組み合わせる必要がある場合といった条件に応じて、参照先を切り替える。これにより、過不足のない情報取得と応答生成を実現し、応答速度の向上や不要なデータアクセスの抑制にも寄与する。

このように、クエリ先選択型 F-RAG を活用した観光案内エージェントにより、単なる定型的な観光情報提供にとどまらず、リアルタイム性と文脈理解を両立した、状況適応型の観光支援サービスを実現する。これは、観光客の満足度向上や回遊促進に寄与すると共に、将来的には防災情報提供や交通案内等、他分野への展開も可能な汎用的アプローチである。

4.2 取得したデータ

(1) 取得したデータ一覧

今回の実証のために機器を設置の上、取得・利用したデータは以下の通りである。

| 項番 | 用途 | データ取得機器名 | 稼働場所 | データ項目 | データ収集期間 | データ形式 |
|----|----------|--|---------------|---|-----------------|--------------|
| 1 | AI 学習 | カメラ(Luxonis OAK-D、WV-U1542LA) エッジボックス(Jetson Orin Nano) | 五稜郭 タワー | 出入口やチケット 売り場、エレベーター前の映像及び 人数カウント値 | 令和7年10 月～11月 | MP4、 JSON |
| 2 | AI 学習 | カメラ(WV-U1542LA) エッジボックス(Jetson Orin Nano) | クルーズ ターミナル | 出入口や建物前 の映像及び人数 カウント値 | 令和7年10 月～11月 | MP4、 JSON |
| 3 | AI 学習 | カメラ(WV-U1542LA) エッジボックス(Jetson Orin Nano) | 函館朝市 エリア | エリアの大通り の映像及び人数 カウント値 | 令和7年10 月～11月 | MP4、 JSON |
| 4 | AI 学習 | カメラ(WV-U1542LA) エッジボックス(Jetson Orin Nano) | 五稜郭公 園前駅 | 市電の乗降場所 や電停の待ち行 列の映像及び人 数カウント値 | 令和7年10 月～11月 | MP4、 JSON |
| 5 | AI 学習 | カメラ(WV-U1542LA) エッジボックス(Jetson Orin Nano) | 函館駅前 駅 | 市電の乗降場所 や電停の待ち行 列の映像及び人 数カウント値 | 令和7年10 月～11月 | MP4、 JSON |
| 6 | AI 学習 | カメラ(WV-U1542LA) エッジボックス(Jetson Orin Nano) | 十字街駅 | 市電の乗降場所 や電停の待ち行 列の映像及び人 数カウント値 | 令和7年10 月～11月 | MP4、 JSON |

(2) データセット作成に係る処理

カメラ由来の人数カウントデータを1分粒度で整形し、品質フラグ及び欠損条件を明確にした上で、混雑予測モデルの学習・評価に用いるデータセットを作成する。データセット作成にあたっては、実運用環境における欠損や推論品質の揺らぎを前提とし、再現性及び比較可能性を確保するため、以下の方針で処理を行う。

- ・各地点の解析結果データを1分粒度の時系列データとして整形する。
- ・地点ID、時刻情報、カメラID等の識別情報を付与し、地点単位で扱える形式に統一する。
- ・学習・評価には、映像そのものではなく、人数・混雑度等の数値データを主に利用する。

- ・逆光、夜間、悪天候、遮蔽等による推論品質低下を考慮し、品質フラグを付与して管理する。
 - ・連続欠損が10分以上発生した区間は、学習・評価対象から除外する。
 - ・外れ値や不自然な急増減が確認された場合は、品質フラグ等を基に採用可否を判定する。
- また、学習データセットとして採用するデータは、以下の条件を満たすものとした。
- ・タイムスタンプが整合し、時系列の連続性が確保されていること
 - ・連続欠損が10分以上の区間を含まないこと(10分以上は除外)
 - ・推論品質フラグが正常であり、遮蔽・極端な逆光・夜間等の低信頼区間でないこと
 - ・外れ値判定に該当しないこと(または補正により妥当性が確保されること)
 - ・個人識別可能な情報、または機密性が高い情報が含まれない粒度に加工されていること

(3) システム全体のデータフロー図

本実証で実施した人流データの取得から分析、情報提供までの流れは図4.2-1の通りである。

現地では、観光地や交通拠点に設置したカメラにより人の動きを把握し、その映像を現地の機器で解析することで、人数や混雑状況といった数値データを取得している。映像データ自体は外部に送信せず、数値情報のみを扱う構成とすることで、通信量を抑えると共に、プライバシーへの配慮を行っている。

取得した人流データはクラウド上で蓄積・分析され、過去の観光統計やイベント情報、天候情報と併せて活用される。これにより、現在の混雑状況だけでなく、今後の混雑の見込みを把握することが可能となる。情報提供においては、利用者の問い合わせに応じて必要なデータを組み合わせ、観光アプリ等の画面を通じて分かりやすく提示する。これにより、観光客は混雑を避けた行動を選びやすくなり、行政や事業者にとっても、混雑緩和や運営判断に役立つ情報を得ることができる。

本構成は、通信環境や設置条件に制約のある地域においても導入しやすく、低コストで継続的な運用が可能で、観光分野に加え、交通や防災、都市計画等、幅広い行政施策への応用が期待される。

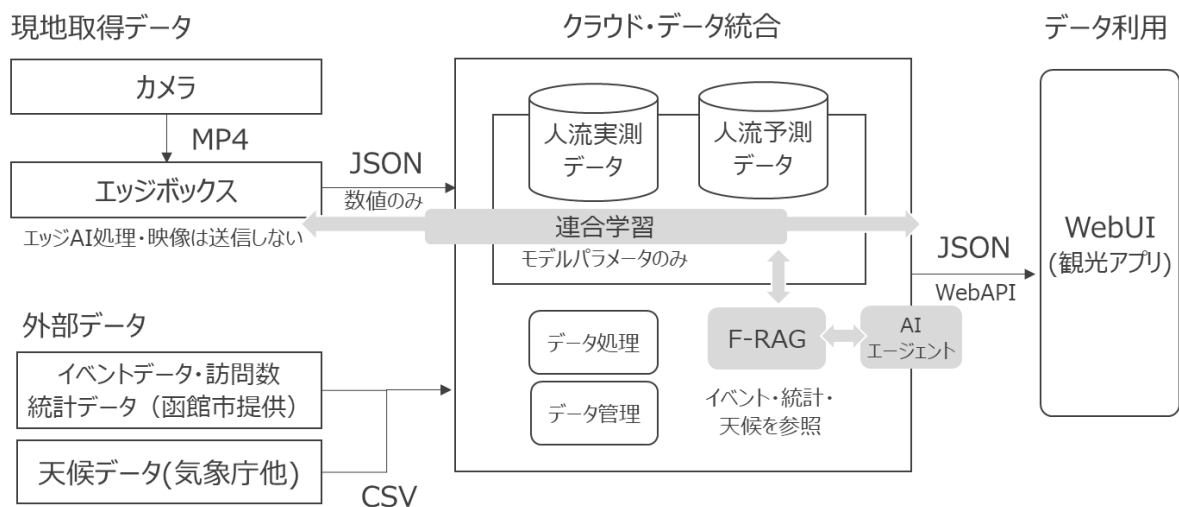


図4.2-1 システム全体のデータフロー図

(4) データの適切さ

本実証で利用したデータは、①カメラによる実測人流データ、②函館市が保有するイベントデータ、③函館市より提供された過去3年分の観光地来訪数に関する統計データ、④気象庁等から取得した天候データの四種を組み合わせたものである。これにより、短期的な人流の実態把握から、中長期的な傾向分析、外部要因(イベント・天候)の影響考慮までを一体的に扱える構成としている。

■データの代表性

カメラによる人流データは、五稜郭タワー、クルーズターミナル、函館朝市エリア、五稜郭公園前駅、函館駅前駅、十字街駅といった、函館市における主要な観光拠点及び交通結節点を対象として取得した。これらの地点は、観光客の滞留・移動が集中しやすく、観光需要や混雑の影響が顕在化しやすい場所であるため、観光・交通分野の混雑把握を目的とした実証として高い代表性を有している。取得期間は令和7年10月から11月までの約2か月間であり、年間全体を網羅するものではないが、秋季の観光需要期を含み、平日・休日差、イベント有無、天候差といった人流変動要因を一定程度含む期間設定となっている。

■データ内容及び品質管理の適正性

カメラは、出入口、チケット売り場、エレベーター前、電停の乗降場所や待ち行列、大通り等、人数の増減や滞留が明確に観測できる位置に設置し、映像データ(MP4)及び解析結果としての人数カウント値(JSON)を取得した。解析及び学習においては、映像そのものではなく人数・混雑度等の数値データを主に利用しており、個人識別を行わない設計とすることで、プライバシー及び利用適正性を確保している。

また、逆光、降雨、夜間等により推論精度が低下する可能性を考慮し、品質フラグの付与及び欠損処理ルール(連続欠損が一定時間以上の場合は学習・評価対象から除外)を設定した。これにより、短期間データであっても品質のばらつきを前提とした適正な学習・評価運用が可能となっている。

■天候データ・イベントデータとの組み合わせによる妥当性

天候データ(天気、降雨・降雪、気温等)及び函館市のイベントデータは、人流に大きな影響を与える外部要因であるため、カメラデータと併せて活用した。これにより、局所的・短期的な人流変動を、天候条件やイベント有無と関連付けて解釈することが可能となっている。

更に、過去3年分の観光地来訪数に関する統計データを併用することで、取得期間が約2か月間であることによる季節的な偏りを補正し、例年傾向との比較という形で分析を行っている。こうして、短期間の実測データを用いた実証において、代表性を補強する。

以上より、本実証で利用したデータは、①観光・交通の主要拠点を対象としている点、②秋季観光需要期における実測データを取得している点、③実測データに行政統計・イベント・天候データを組み合わせる点から、取得期間が約2か月間であるという制約を踏まえても、代表性及び適正性の観点から妥当なデータ構成であると評価できる。

4.3 AIモデルの学習プロセス

本実証では、連合学習(Federated Learning)方式に基づき、各拠点での学習(ローカルトレーニング)を前提とするため、実証期間中の学習・検証は実際のエッジ端末総統のローカル環境を用いて実施した。学習対象は、人数時系列から近未来を推定する予測モデルである。学習はローカル更新として複数回試行し、更新パラメータを集約してグローバルモデルを構成するプロセス(FedAvg 等)を実施する。

(1)学習環境(ローカルトレーニング実施環境)

- OS:Windows 11
- CPU:8~コア級(Intel Core i7 12 世代)
- GPU:NVIDIA RTX 3060 クラス(VRAM 12GB 程度)
- メモリ:32GB
- ストレージ:NVMe SSD 1TB
- 主要ライブラリ:PyTorch / TensorFlow、Flower/FedML 等の FL フレームワーク

(2)学習手順(ローカル学習→集約→評価)

- 1.各拠点データ(または実証データ)を用いてローカルモデルを初期化
 - 2.ローカル環境でミニバッチ学習(数エポック)を実行し、モデル更新量(重み差分等)を生成
 - 3.更新量をサーバ側で集約(FedAvg 等)しグローバルモデルを更新
 - 4.更新後のグローバルモデルを各拠点へ配布(もしくは評価用に適用)
 - 5.上記を所定ラウンド回(例:10~50 ラウンド)繰り返し、性能・収束性を確認
- 検証データに対する RMSE で性能評価の上、以下(i)及び/または(ii)をもって収束性を判断する。
- (i)グローバルモデルの検証 RMSE の改善量が所定の閾値以下となる状態が連続して続くこと
(例:3~5 ラウンド連続で $\Delta\text{RMSE} < \epsilon$)
 - (ii)連続ラウンドにおけるモデル更新量(例: $\|\Delta w\|$)が十分小さくなること
- また、検証 RMSE が早期に頭打ちや悪化に転じる等、停滞や局所解が疑われる場合には、複数の初期値(乱数 seed)で学習を再実行し、最良の検証 RMSE および安定性(ラウンド間変動が小さいこと)を示すモデルを採用する。必要に応じて学習率、ローカルエポック数、クライアント選択率等のハイパーパラメータも再調整し、同条件で再評価する。

•推論環境(エッジ推論/グローバル推論)の想定スペック

(A)エッジ推論(現地設置を想定したデバイス)

- 現地での推論は「映像を外部へ出さずに人数・混雑度を算出する」ことを目的とするため、省電力で常時稼働可能なエッジ機器を想定する。推論は 1~5 FPS 程度(混雑把握用途)を想定し、以下のいずれかの構成を推奨とする。
- 構成(GPU/AI アクセラレータ搭載のエッジ)
デバイス:NVIDIA Jetson Orin Nano
CPU:6~8 コア級(ARM)、RAM:8~16GB
推論性能目安:軽量 CNN(Mobile Net/Efficient 系)で 1~10 FPS 程度
想定用途:現地カメラ入力→人数推定→混雑度(数値)出力→ローカル保存

・エッジ推論では、画像は端末内で処理し、外部送信は「人数・混雑度」等の数値データに限定する設計とした。

(B) グローバルモデル推論(サーバ/クラウド/拠点サーバ)

・グローバルモデル(集約後モデル)の推論は、混雑予測や観光案内(混雑回避提案等)に利用することを想定し、時系列モデル推論が中心となる。時系列モデルは一般に画像モデルより軽量であるため、サーバ側は CPU 中心でも運用可能だが、複数地点・高頻度リクエストに備えた余裕を見込む。

・構成(中規模サーバ/クラウド VM)

CPU: 8~16 vCPU、RAM: 32~64GB

GPU: 必須ではない(画像推論も集約して行う場合は T4/A10 相当を追加)

推論性能目安: 多数地点の時系列予測(例: 数十~数百地点×数分更新)に対応可能

用途: 集約済みモデルを用いた予測 API

・学習の所要時間・試行回数

・「PC(RTX 306 級)」を前提とした測定値

・画像系(混雑度推定)モデル:

学習時間: 1 試行あたり 30 分~3 時間程度(データ量・解像度・モデル規模による)

試行回数: ハイパーパラメータ調整等を含め 3~10 回程度

時系列予測モデル:

学習時間: 1 試行あたり 5 分~60 分程度(系列長・特徴量数・モデル規模による)

試行回数: 特徴量追加・予測ホライズン変更等を含め 5~20 回程度

連合学習ラウンド:

1 ラウンド(ローカル数エポック+集約)の所要: 数十秒~数分

ラウンド回数: 10~50 ラウンド程度(収束確認を含む)

4.4 AI モデルの開発・学習にあたっての留意事項等

本実証における AI モデル開発及び学習では、観光地という実環境特有の制約(通信条件、設置制約、プライバシー配慮、季節変動、イベント影響等)を踏まえ、以下の点を重視して重視して設計・運用を行った。以下に、主な工夫点及び運用上の知見を整理する。

■工夫した事項

(1) 「映像を出さない」を前提とした二段階設計(画像→数値→予測)

画像はエッジで推論し、拠点外に出さず、学習・共有の中心は人数・混雑度等の数値時系列とした。これにより、プライバシー・機密性の懸念を抑えつつ、通信負荷も低減し、導入可能地点を拡張できる。Tips: まずは「人数(数値)」の品質を上げる(推論の安定化)ことが、予測精度向上の近道となる。

尚、人物検出には YOLO 系アルゴリズムをベースとしたモデルを採用したが、本エッジ AI は、人物の全身、頭部、顔といった複数の検出方式を組み合わせるモジュールとして実装している。予測精度向上に向けて、以下を取り組むことで、入力データ品質の安定化と設置環境に応じた検出方式の最適化を実施した。

・設置環境に応じた検出方式の切り替え:

全身・頭部・顔検出を設置条件や混雑状況に応じて使い分けることで、誤検知や欠損を抑え、人数カウント精度のばらつきを低減した。

・流入・流出を含む人流構造の把握：

エリア滞留人数に加え、クロスライン・クロスエリアカウントを実施することで、単なる人数ではなく人の動きの変化を捉え、混雑の増減傾向をモデルに反映できた。

(2) 実運用を想定したデータ品質フラグ運用

遮蔽・逆光・夜間等、推論品質が低下する条件を「品質フラグ」として管理し、学習採用条件・評価条件に反映した。

Tips:学習前のクレンジングよりも、「低品質区間を明示的に除外・別扱い」する方が、現場運用と整合しやすい。

(3) 欠損処理のルールを先に固定(再現性・比較可能性)

連続欠損 10 分以上は除外という採用条件を明確化し、欠損補完の恣意性を排除した。

Tips:欠損条件は事後に変更すると精度が動きすぎるため、最初に固定しモデル改善は別軸で行う。

■計画から差分が生じた事項

(1)当初想定より「現地データの品質揺らぎ」が大きかった

要因:観光地は天候・季節・イベント・時間帯で照度や人流が大きく変動し、カメラの視野遮蔽・逆光・夜間等が発生しやすい。これにより人数推定の誤差・欠損が増え、学習にそのまま投入すると予測モデルが不安定化する。

対処法:品質フラグ(遮蔽・低照度等)の導入、外れ値除外、欠損処理ルールの固定(連続欠損 10 分以上除外)、必要に応じて平滑化を実施する。

(2)予測が難しい「急激な増減(スパイク)」への対応が必要になった

要因:団体バス到着等による流入スパイクは、過去時系列だけでは説明できず、予測誤差が残りやすい。

対処法:外部要因(天候・曜日・イベント等)の特徴量追加、上流地点(駐車場・バス降車地点等)の観測追加／連携検討、複数地点の相互作用を扱うモデル構造を検討する。

影響:モデル構造・入力設計の見直しが発生し、一部評価指標の再定義(通常時精度とイベント時精度の分離)が必要となった。一方で、実運用に耐える改善方針が明確化した。

(3)連合学習の実装は「学習」よりも「運用設計(参加条件・更新頻度)」が支配的になった

要因:拠点毎にデータ量・欠損率・稼働時間が異なり、同一条件での学習参加が難しい。更に通信状況により更新が失敗する場合がある。

対処法:参加条件(一定期間のデータ量、欠損率閾値等)を定め、更新頻度を抑えた運用(例:夜間バッチ更新)を優先する。まず時系列モデルから連合学習の安定運用を検証する。

影響:当初想定より連合学習ラウンド数・更新スケジュールの設計に工数が必要となったが、結果として「現場で回る連合学習」の要件整理ができ、他地域への移植性が高まった。

■他地域で同様のモデル構築を行う際の Tips

- ・設置前に“品質が落ちる条件”を洗い出す(逆光・夜間・遮蔽・降雪等)。品質フラグを最初から設計に入れる。
- ・欠損ルールを固定してからモデル改善に入る(例:連続欠損 10 分以上除外)。
- ・連合学習は「小さいモデルから回す」(時系列予測モデル中心→必要に応じて画像モデルへ拡張)。
- ・通常時とイベント時を同じ指標で評価しない(通常時 MAE + イベント時のスパイク再現率等、二系統評価が実務的)。
- ・現地で合意形成が難しいデータほど“連合学習前提”が効く(データを外に出さずに価値を得られる構造をつくる)。

5. 実証の手法

5.1 ユースケース①:エッジ AI による通信量の低減

当該ユースケースは実施しなかった。

5.2 ユースケース②:山間部や海中等の携帯電話網不感地域における通信の確立

当該ユースケースは実施しなかった。

5.3 ユースケース③:多種多様なデバイスに対する通信リソースの最適化

当該ユースケースは実施しなかった。

5.4 ユースケース④:分散ネットワークにおける新たな AI ソリューションの創出

5.4.1 実証内容の詳細

各観光エリアに設置するエッジデバイスで取得された人流データを、エッジサーバを用いた連合学習により予測モデル生成する。また、この結果を、AI エージェントが、エリア特有のパターンを考慮して効率的に収集し、混雑予測を実施する。更に、この結果を活用して、大規模言語モデルを用いた観光案内エージェントを構築する。この際、分散型ネットワーク上において、連合学習(Federated Learning)やエッジ AI を活用することで、クラウドだけに依存せずに、高精度な AI 処理を実現する仕組みを構築する。観光や交通分野において、プライバシーを保ちつつリアルタイムな混雑予測やデータ統合を可能にする「軽量・効率型の AI ソリューション」を構築することが本ソリューションの特徴である。(図 5.4-1)

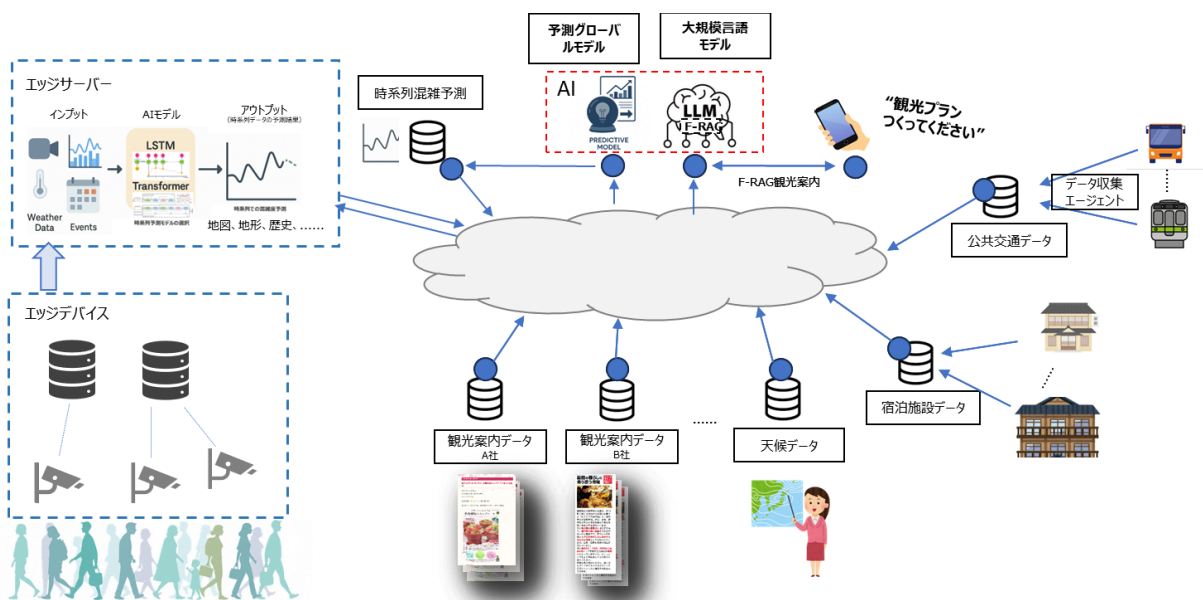


図 5.4-1 AI 技術設計 全体像

本ソリューションの特徴は、通信負荷を最小限に抑えながらも、高精度な混雑予測の実現を両立することである。この点について検証を行うため、本実証では、函館市内において、観光地や交通駅等の環境が異なる観光スポットに設置したカメラから取得した人流データについてエッジ AI を用いた混雑予測システムを構築し、KPI として設定している通信コスト削減率や予測精度等を確認することにより、通信量の削減と混雑予測の精度の両立の実現性を検証する。

5.4.2 利用技術・データ

まずデータ取得における AI 技術として、エッジデバイス上での映像解析を行うため、YOLO 系モデルをベースとした人物検出 AI を採用した。本モデルは、全身・頭部・顔といった複数の検出方式を使い分けることが可能であり、カメラの設置高さや角度、混雑度合いに応じて検出精度を最適化できる点が

特徴である。これにより、人物同士の重なりが発生しやすい観光地においても、安定した人流データ取得が可能となっている。

取得された人流データは、時系列データとして数値化・集計の処理が実施され、函館市から提供された観光イベント情報や過去の来訪傾向データ、並びに天候データと組み合わせた上で、LSTM や Transformer を用いた時系列予測モデルにより混雑予測に活用される。これらのモデルは、過去の混雑傾向だけでなく、天候、曜日、時間帯、イベント情報等の外生要因を考慮できる点に特徴があり、突発的な混雑や季節変動を含めた動的な予測を可能としている。この混雑予測 AI は、観光客への事前案内や回遊促進施策を検討する上で重要な役割を担う。

更に、本実証では、連合学習(Federated Learning)を採用し、各観光エリアで取得された人流データをクラウドに集約することなく、モデルパラメータのみを共有・統合する構成とした。Federated Averaging(FedAvg)を用いることで、地域毎の特性を反映したローカルモデルを維持しつつ、全体として高精度なグローバルモデルを構築できる点が大きな特徴である。この手法は、個人情報や映像データを外部に送信しないという点で、観光・交通分野における AI 活用の社会受容性を高めるうえで極めて重要である。

また、生成された混雑予測結果については、AI エージェントが分散環境下で効率的に収集・統合し、必要なデータのみを選択的に判断の上、活用する。これにより、通信量や処理負荷を抑えつつ、リアルタイム性の高い混雑予測を実現している。

更に、観光案内への応用としては、大規模言語モデルとクエリ先選択型 F-RAG を組み合わせ、一般的な観光情報に加え、リアルタイムの混雑予測や交通状況を踏まえた自然言語での情報提供を可能とした。この構成により、単なる定型情報ではなく、利用者の状況や質問意図に応じた柔軟な案内が実現されている。

これらの各処理フローについて、データ取得から利活用までの一連のフェーズとして整理すると以下の表 5.4-1 の通りとなる。

表 5.4-1 データ取得から活用フェーズにおける利用技術・データ

| フェーズ | 利用技術 | 利用データ |
|-------|------------------|-------------------------------|
| データ取得 | エッジ AI(CNN) | 映像 → 人数カウント |
| 前処理 | エッジ処理 | 数値化・集計 |
| 予測 | LSTM/Transformer | 人流+天候+イベント (イベント・統計:函館市提供) |
| 学習統合 | 連合学習(FedAvg) | モデルパラメータ |
| 判断 | AI エージェント | 各エリア予測結果 |
| 提供 | LLM+F-RAG | 混雑予測+観光情報 (観光情報:函館市提供) |

5.4.3 必要性・新規性

■実施内容の必要性

①観光・交通分野における構造的課題

函館市はオーバーツーリズムの問題を抱えており、少子高齢化の課題を抱えている。インバウンドを含む観光客の増加や大型客船の到着により、公共交通機関の運転手や観光地での現地スタッフの不足等の課題を抱えており、観光需要のリアルタイムな状況把握や人員・車両配置計画の策定にとって、今回の取り組みに対して提供ニーズが極めて高い。以上のようにオーバーツーリズムや時間帯・季節による混雑偏在が顕在化している一方で、リアルタイム性と将来予測の双方を満たす混雑把握手段が不在している課題がある。これらの課題は地域都市に共通しているため、今回の実証として相応しいと考える。

②社会実装を前提とした必要性

観光分野における AI・データ活用は、単なる実証にとどまらず、住民・事業者の受容性や長期的な運用コスト、他地域・他分野への展開可能性を考慮しなければ社会実装に至らない。本実施内容は、初期投資・通信負荷・運用負荷を抑えつつ、持続的に使える構成を前提としており、この点においても必要性が高い。

■新規性(既存取組との差分)

本実施内容は、エッジ AI、連合学習、AI エージェント、LLM を統合することで、低通信負荷・高精度・プライバシー配慮を同時に実現する分散型 AI 基盤を構築しており、既存のクラウド集中型や単点 AI 活用の取組とは明確に異なる新規性と市場優位性を有している。

①市場における既存の主な取組み

既存の取組としては、主に表 5.4-2 のような事例があり、いずれも一定の有用性を有するものの、リアルタイム性、予測性、プライバシー配慮、通信効率といった観点を同時に満たすことが難しいという共通の課題を抱えている。これらの限界を踏まえ、本実証では分散型 AI インフラを用いた新たなアプローチの有効性を検証することとした。

表 5.4-2 既存の取組における主な特徴と課題

| 類型 | 主な特徴 | 限界・課題 |
|-------------|----------------------|--------------------|
| クラウド集中型映像解析 | 映像をクラウドに集約し AI 解析 | 通信コスト大、遅延、プライバシー懸念 |
| センサー・統計ベース | Wi-Fi/Bluetooth、過去統計 | 精度・リアルタイム性が限定的 |
| 単点 AI カメラ | 単一拠点での人数カウント | エリア横断・予測が困難 |
| 観光アプリ型情報提供 | 静的情報・混雑目安 | リアルタイム性・予測性に乏しい |

②本実施内容の新規性・差別化ポイント

・エッジ AI × 連合学習による「分散・高精度」混雑予測

既存の混雑解析システムでは、カメラ映像や人流データをクラウド上に集約し、集中処理によって解析・予測を行う方式が一般的である。一方、本実施では、エッジデバイス上で映像解析及びローカル学習を行い、学習済みモデルのパラメータのみを共有・統合する連合学習方式を採用している。この構成により、映像データや個人に紐づき得る情報を外部に送信することなく、モデル精度を継続的に向上させることが可能となり、プライバシーへの配慮と高精度な混雑予測を両立している。また、データは各観光エリアや交通拠点に留めたまま活用されるため、データの管理権限や利用範囲を地域・運用主体が保持できる「データ主権」を担保した設計となっている。更に、各観光エリアや交通拠点の特性をローカル学習により反映できるため、地域特性に適応した予測モデルを構築できる点も大きな特徴である。このようなデータ主権・プライバシー保護を前提とした連合学習を、実運用レベルで人流・混雑予測に適用する事例は市場においても限定的であり、本実施の重要な新規性の一つである。

・通信前提を変える「軽量・効率型 AI 基盤」

既存の AI 活用型観光・人流分析システムの多くは、高帯域かつ安定した通信環境を前提とし、映像や大量データをクラウドへ送信する構成となっている。これに対し、本実施では、エッジ AI による前処理と数値データ化を前提とし、低帯域(kbps レベル)の通信環境でも成立する設計を採用している。これにより、クラウドへのデータ集中を避けるだけでなく、通信経路上を流れる情報自体を最小限かつ非個人情報に限定することで、セキュリティリスクの低減にも寄与している。加えて、MQTTS(TLS)やVPNを用いた暗号化通信を前提とすることで、モバイル回線利用時や公衆網を経由する環境下においても、安全性を確保している。その結果、モバイル回線を用いた運用や、災害時・通信制約下においても継続的なデータ収集・解析が可能となり、通信費用やインフラ依存度を大幅に低減できる。特定の通信事業者や高性能ネットワークに依存しない点は、セキュリティ・レジリエンスの観点からも新規性が高く、地方都市や他地域への展開における再現性の高さにつながっている。

・AI エージェントによる「自律的・選択的データ活用」

既存システムでは、取得可能なデータを常時収集・処理する設計が一般的であり、通信量や処理負荷が過剰となるだけでなく、不要なデータアクセスがプライバシー・セキュリティ上のリスクとなるケースも多い。本実施では、AI エージェントを用いて、必要なデータのみを、必要なタイミングで、必要な範囲に限定して取得・統合する仕組みを導入している。これにより、通信量や計算資源の利用を動的にかつ最小限に最適化できる。観光エリア毎の特性や時間帯、イベント発生時等の状況変化にも即応でき、かつ不必要なデータ流通を抑制できる点は、分散環境におけるセキュリティ設計としても先進的である。このような自律的・選択的データ活用を前提としたアーキテクチャは、従来の集中型システムとは異なる設計思想であり、本実施の新規性を構成する重要な要素である。

・混雑予測 × LLM × F-RAG による「状況適応型観光案内」

既存の観光案内サービスは、静的な観光情報の提示や、簡易的な混雑目安の表示にとどまるものが多く、またデータを一元的に集約する構成が一般的である。本実施では、混雑予測結果と大規模言語モデル(LLM)を組み合わせ、更にクエリ先選択型 F-RAG を用いることで、分散管理されたデータを安全に参照する構成を採用している。F-RAG により、LLM はユーザーの質問内容に応じて、必要なデータソースのみを選択的に参照し、元データを集約・複製することなく回答を生成する。このため、混雑予

測データや交通データの管理主体・更新権限を維持したまま、リアルタイム性と文脈理解を両立した観光案内が可能となる。これにより、「現在及び今後の混雑状況」を踏まえた行動提案や、空いている観光地・移動ルートの案内が可能となり、観光回遊の促進や混雑分散への直接的な貢献が期待できると同時に、データ主権・プライバシー・セキュリティを損なわない形で高度な観光案内を実現している点が、本実施の大きな差別化ポイントである。

5.4.4 検証条件

①地理的・空間的条件

本実証は、北海道函館市内の観光エリア及び公共交通結節点を対象として実施した。対象エリアは、市内でも特に観光需要が集中する地域であり、歴史的観光資源、商業施設、公共交通機関(路面電車)が比較的限られたエリア内に集約されているという特徴を有する。具体的には以下のような人流が集中・分散する代表的な空間特性を有する地点を対象とした。

- ・観光客の滞留が発生しやすい主要観光スポット
- ・観光スポット間の移動動線上に位置する市街地エリア
- ・路面電車の主要電停(観光地最寄り駅)

これらの場所はいずれも、市街地に位置し、歩行者・車両・公共交通が混在する実環境であり、実運用を前提とした AI インフラ検証に適した空間条件となっている(実際の設置場所は図 3.2-1 の通り)。

②時間的条件

本実証は、観光需要が一定程度存在する期間として 10 月から 11 月までの約 2 か月間において実施した。実証期間中は、以下のような観光客及び市民の行動パターンが変化する時間帯を継続的に観測した。

- ・平日・休日の双方を含む
- ・昼間時間帯を中心としつつ、朝夕の移動ピークを含む

また、短期的なイベント日と通常日が混在する期間であり、日別・時間帯別に人流の変動が発生する条件下でデータ取得及び解析を行っている。これにより、単一条件に依存しない、実運用に近い時間的変動を含む環境を前提とした実証となっている。

③天候・自然環境条件

実証は、北海道函館市の気象条件下で実施しており、晴天・曇天・降雨・低温環境といった、複数の天候・環境条件が含まれている。特に函館市は、季節や天候による観光行動の変化が大きい地域であり、天候が人流や混雑状況に与える影響を無視できない環境条件である。また、屋外設置機器については、風雪・低温・降雨・湿度といった自然環境条件にさらされることを前提としており、これらの条件下でも継続的にデータ取得・通信・解析が可能な構成で実証を実施した。

5.4.5 開発・評価項目

| 番号 | 開発・評価項目 |
|----|--|
| 1 | データ環境の整備 AI エージェントが巡回時に閲覧可能なデータ環境を構築(当初は擬似環境) |
| 2 | エージェントの開発 巡回アルゴリズムを開発し、エッジやクラウドの情報へアクセスできること |
| 3 | 分散連合型 RAG・エージェント RAG の実装 分散 RAG の環境を準備し、混雑予測情報を取得し LLM が正しい回答を実施できること |
| 4 | 連合学習の実装 エッジボックス上でその場の混雑データを学習して、モデルのみクラウドで統合できること |
| 5 | エージェントとリアルワールドデータの連携 連合学習により生成されたモデルを利用して分散 RAG 上の混雑データを AI エージェントが巡回して取得し混雑予測 AI が観光地の近い未来の混雑予測結果を推論できること |
| 6 | 実証実験 実際のデータを利用して、KPI・KGI の検証を実施の上、AI エージェントの巡回アルゴリズムの改善や混雑予測を実施し、データ同化により予測精度の改善を繰り返す。また連合学習の結果が、全データ集約と比較し、通信負荷・リアルタイム性・プライバシー保護、適応性に問題がないか検証する。 |

1) データ環境の整備

カメラ映像データ並びに天候・曜日・時間帯等の外部データについて、学習用データセット作成のために、プライバシー保護及び学習品質の担保の観点から、以下の処理を実施する。尚、生データ(画像等)は原則として拠点外へ持ち出さず、学習に用いるデータは人数・混雑度等の数値化データ(時系列)を主とする。

①データ整形・統合(前処理)

- ・時刻基準の統一:タイムスタンプを統一し、重複・ずれ・不整合を補正する。
- ・サンプリング間隔の統一:推論頻度(例:1分、5分等)に合わせ、人数・混雑度の時系列を一定間隔にリサンプリングする。
- ・外部データの付与:天候、曜日、祝日、イベント有無、時間帯等の外生特徴量を同一タイムスタンプに紐づけ、学習入力形式へ整形する。

②欠損・異常値・ノイズへの対応(品質管理)

- ・欠損値処理(採用条件を明確化):
連続欠損が10分未満の短時間欠損については、前値保持や線形補間等により補完する。
連続欠損が10分以上の区間は、センサー・通信・推論処理の停止等により系列の信頼性が担保できないため、学習データセットから除外する。
- ・異常値(外れ値)除外:推論誤検出等に起因する不自然な急増減(例:急激なゼロ化、物理的に不可能な値)を検知し、除外または補正する。

・低品質区間の除外: 遮蔽・逆光・夜間等、推論精度が著しく低下する条件下のデータには品質フラグを付与し、学習採用条件から除外する。

・ノイズ低減: 必要に応じて移動平均等の平滑化処理を行い、短周期ノイズの影響を低減する。

③プライバシー・機微情報保護のための加工・除外

・連合学習前提の取り扱い: 生データは拠点内に保持し、拠点外へはモデル更新情報のみを送受信する運用を前提とする。

2) エージェントの開発

AI エージェントの開発にあたっては、OpenAI が提供する大規模言語モデル(LLM)API を活用し、対話処理及び推論制御を行うエージェントを構築する。本実証では、LLM を単体で利用するのではなく、外部データ取得・推論結果の統合・応答生成を行うエージェント機能の中核として組み込む。

3) クエリ先選択型 F-RAG の実装

複数のデータソースの中から選択的に情報を取得し、回答を行うクエリ先選択型 F-RAG を実装し、混雑予測情報が正しく情報提供できているかを確認する。

具体的なアーキテクチャは以下の図 5.4-2 の通りである。LLM に OpenAI 社の GPT を利用し、MCP(Model Context Protocol)と呼ばれるプロトコルを通じて、AI エージェントが各種データソースから必要な情報を取得できるようにする。

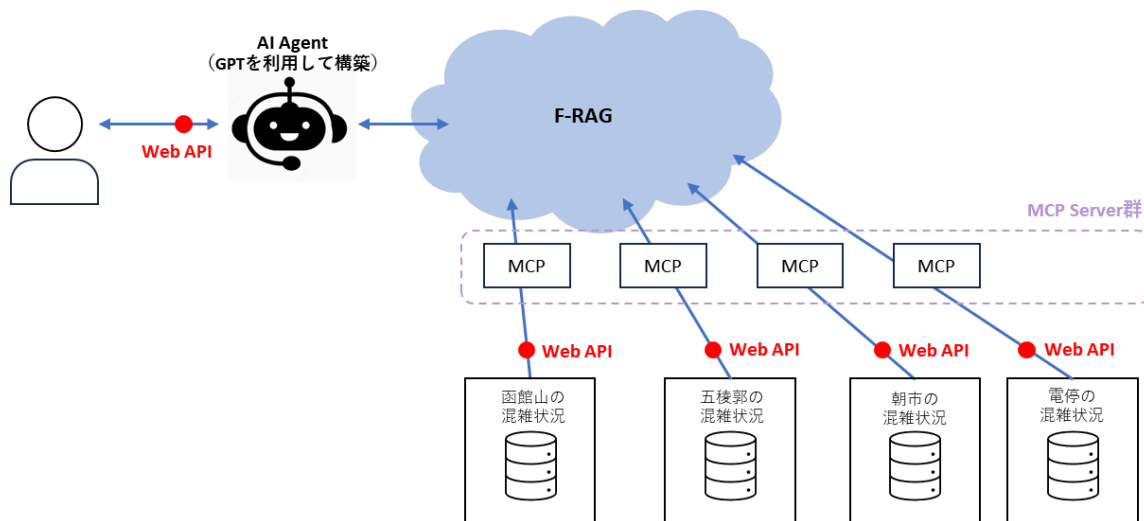


図 5.4-2 F-RAG のアーキテクチャ

本実証においては、WebAPI から取得される混雑予測のデータ等について MCP を通じて AI Agent が取得でき、かつ、それをユーザーに提供できるかを確認する。この他、本システムを用いて、本実証で生成する混雑予測結果やその他データを取得してユーザーに提示する観光案内エージェントを作成する。

4) 連合学習の実装

本モデルのアーキテクチャは、予測対象となる課題の特性を考慮して戦略的に選定されている。特に、個々の地点における混雑度の時間的推移と、複数の観光地間に存在する空間的な相互依存関係を同時に捉えるため、ハイブリッドアプローチを採用した。この設計により、単一の時系列分析では捉えきれない、より複雑で動的な観光客の動態パターンを学習することが可能となる。

■ハイブリッドアーキテクチャの概要

モデルの核となるのは、時間的特徴と空間的關係性を学習する 2 つの主要コンポーネントを組み合わせたハイブリッドアーキテクチャである。具体的には、Temporal Convolutional Network (TCN) と Node-wise Self-Attention 機構で構成されている(図 5.4-3 の通り)。

・TCN (Temporal Convolutional Network):

各観測地点(ノード)の来訪者数のような時系列データ内に存在する、長期的なパターンや依存関係を効率的に学習する役割を担う。

・Node-wise Self-Attention:

複数の観光地をグラフのノードとみなし、それらのノード間で動的に変化する相互作用や影響度を学習する。これにより、ある地点の混雑が他の地点に与える影響等をモデルに組み込むことができる。

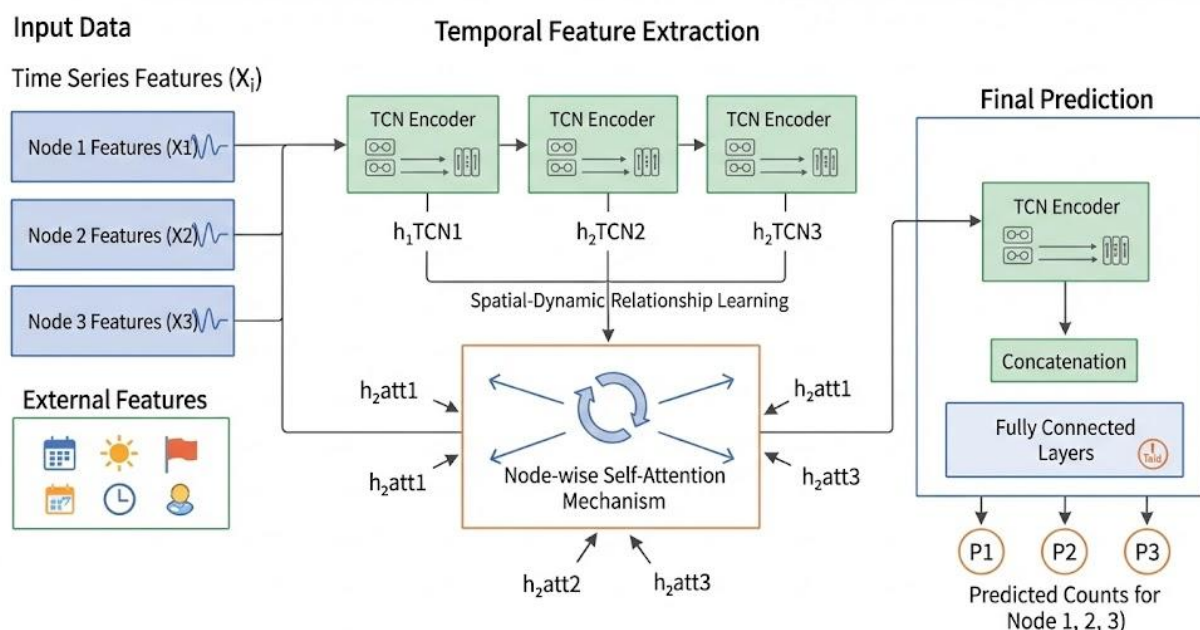


図 5.4-3 ハイブリッドアーキテクチャ概要図

本構成は、各地点の時間的変化と、地点間の相互影響を同時に学習することで、単一地点では捉えきれない動的な人流・混雑パターンを高精度に予測するモデルである(「時系列・空間統合型混雑予測モデル」)。尚、各用語の定義については、表 5.4-3 の通り。

表 5.4-3 時系列・空間統合型混雑予測モデル 用語定義一覧

| 区分 | 用語 | 定義・説明 |
|---------|---------------------------------------|--|
| 入力データ | Input Data | モデルに入力される全データ群。混雑予測の基礎情報となる |
| | Time Series Features(X_i) | 各ノード(観光地・地点)の過去の時系列データ 人数カウントや混雑度など、地点固有の時間変化を表す |
| | Node(ノード) | 予測対象となる地点単位(観光スポット、カメラ設置地点等) |
| | External Features (外部特徴量) | 全ノードに共通して影響する外部要因データ。天候、曜日、時間帯、イベント有無等 |
| 時間的特徴抽出 | Temporal Feature Extraction | 時系列データから時間的パターンを抽出する処理全体 |
| | TCN Encoder | Temporal Convolutional Network:長期的な時系列依存関係を効率的に学習するニューラルネットワーク |
| | h_1 TCN1 / h_2 TCN2 / h_3 TCN3 | 各ノードに対して TCN Encoder が出力する時間的特徴表現 |
| 空間的関係学習 | Spatial-Dynamic Relationship Learning | ノード間の影響関係(人流の伝播等)を学習する処理 |
| | Node-wise Self-Attention Mechanism | 各ノード間の影響度を動的に算出し、空間的相関を学習する仕組み |
| | Self-Attention | ノード同士の関係性に重み付けを行い、相互作用を表現する注意機構 |
| | h_2 att1 / h_2 att2 / h_2 att3 | Attention 処理後に得られる、空間的相互影響を反映した特徴表現 |
| 最終予測処理 | Final Prediction | 抽出した特徴を統合し、最終的な予測値を算出する処理。 |
| | Upper TCN Encoder | Attention 後の特徴をさらに統合し、予測向けに高次特徴を抽出する TCN 層 |
| | Concatenation | 複数の特徴量を結合し、一つの特徴ベクトルにまとめる処理 |
| | Fully Connected Layers | 結合された特徴を最終的な数値予測へ変換する全結合層 |
| 出力 | Predicted Counts (P1, P2, P3) | 各ノードに対する将来時点の混雑人数・混雑度の予測結果 |

■エッジにおけるデータ処理フロー

モデル内部におけるデータ処理は、以下のステップで実行される(機械学習モデルの詳細な構成は付録3にて記載の通り)。

1.入力 (Input): 各観測地点(Node)の時系列データ(X_i)と、天候や曜日といった外部特徴量がモデ

ルに入力される。

2.時間的特徴抽出 (Temporal Feature Extraction): 各ノードの時系列データは、一連の TCN エンコーダに入力される。ここでデータ内の時間的なパターンが抽出され、時間的特徴量(h_t ,TCN)が生成される。

3.空間的・動的関係学習 (Spatial-Dynamic Relationship Learning): TCN によって抽出された各ノードの時間的特徴量は、次に Node-wise Self-Attention 機構に渡される。この機構がノード間の動的な関連性を学習し、他のノードからの影響を考慮したアテンション加重特徴量(h_t ,att)を生成する。

4.最終予測 (Final Prediction): アテンション加重特徴量は、最終段の TCN エンコーダで更に処理された後、連結(Concatenation)され、全結合層(Fully Connected Layers)を通過する。これにより、各ノードの将来の来訪者数(P1, P2, P3)が最終的な予測値として出力される。

■連合学習における処理フロー(Federated Averaging)

TCN を用いた連合学習は、以下のループで実行される。

1.初期化: サーバが TCN の初期モデル(重みランダム、または事前学習済み)を作成する。

2.Broadcast: サーバが全エッジ(五稜郭、駅、空港等)にモデルを配布する。

3.Local Update (エッジ処理):

各エッジが手元のデータ(直近 1 週間分等)を使って TCN を数エポック学習する。

「自分の場所の傾向」をモデルに反映させる。

4.Upload: 各エッジが「更新された重みパラメータ」だけをサーバに送信する。

5.Aggregation (サーバ処理):

サーバは集まった重みを平均化する(データの件数に応じた重み付け平均)。

$$\text{New_Global_Weight} = \sum (\text{Client_Weight} * \text{Data_Size}) / \text{Total_Data}$$

これにより、「五稜郭のトレンド」と「駅のトレンド」が混ざり合った、汎用的なモデルが完成する。

6.Repeat: 手順 2 に戻る(これを一定期間毎に繰り返す)。

■データ仕様

モデルの予測精度は、その入力データの質と多様性に根本的に依存する。本モデルでは、リアルタイムで観測される時系列データと、状況を補足する外部特徴量を組み合わせることで、精度の高い予測を実現している。入力データとして、モデルの学習及び予測に使用されるデータは、時系列データである過去 L 時間分の各観測地点における人数と外部特徴量である天候、時刻、曜日の 2 つのカテゴリに分類される。

■観測地点(ノード)

予測対象となる観測地点は、以下の通り函館市内の主要な観光エリアに設置されている。

- 五稜郭タワーエリア:
 - ・goryokaku_tower_overall (全体)
 - ・goryokaku_tower_ticket_counter (チケット購入窓口)
 - ・goryokaku_tower_1f_elevator (1 階エレベーター)
 - ・goryokaku_tower_observation_floor (展望台フロア)
- 電停エリア:

- ・goryokakukoenmae_station_alignments (五稜郭公園前駅:降車数)
- ・goryokakukoenmae_station_bordings (五稜郭公園前駅:乗車数)
- ・hakodateekimae_station_alignments (函館駅前駅:降車数)
- ・hakodateekimae_station_bordings (函館駅前駅:乗車数)
- ・jujigai_station_alignments (十字街駅:降車数)
- ・jujigai_station_bordings (十字街駅:乗車数)
- 函館朝市エリア:
 - ・asaichi_overall (全体)
 - ・asaichi_ekimaeparking (駅前駐車場の前エリア)
 - ・asaichi_nakadoori (仲通り側エリア)
 - ・asaichi_kaikoudoori (開港通り側エリア)
 - ・asaichi_parkingmaedoori (朝市駐車場通りエリア)

■出力データ

モデルの出力は、各観測地点(ノード)における H 時間後の予測人数である。これらのデータ仕様に基づき、モデルがどの程度の性能を発揮するかを評価する。

5) エージェントとリアルワールドデータの連携

ウェブインターフェイスからエージェントへのアクセスが可能な環境を構築する。クラウド上に配置した AI エージェントの機能を Web API として提供する。ウェブインターフェイスでは、利用者が混雑状況や混雑予測に関する問い合わせを行うと、その内容が API を通じて AI エージェントへ送信される。AI エージェントは、問い合わせ内容に応じて、分散 RAG (F-RAG) を用いて必要な地点・時間帯の混雑データ(実測値及び予測値)を取得し、大規模言語モデルと組み合わせて応答を生成する。生成された応答結果は、再び API を介してウェブインターフェイスへ返却され、画面上に表示される。こうして、カメラの実データによるリアルタイムな混雑度の可視化に加えて、函館の観光情報、混雑予測、F-RAG による観光案内を組み込む。

6) 実証実験

実データを用いて KPI・KGI を検証し、応答性能、予測精度、通信特性を計測しながら、AI エージェント及び予測モデルの改善を反復的に実施する。実際に上記ウェブインターフェイスを開放し、函館の観光客を想定して、検証する。観光客の行動を想定した問い合わせシナリオを設定し、ウェブインターフェイスを公開する。実際の混雑実測データ及び予測データを用いて、AI エージェントによる応答を取得の上、API から応答時間・通信量・成功率等を計測する。予測結果と実測値を比較し、予測精度及びデータ同化による改善効果を評価する。また、評価結果を基に、AI エージェントの巡回アルゴリズムや予測モデルを改善の上、連合学習方式と全データ集約方式を比較し、通信負荷、リアルタイム性、プライバシー保護、適応性の観点から問題がないかを確認する。

KPI・KGI との対応を明確化するため、評価観点を①通信負荷低減、②AI 性能、③混雑予測公開による行動変容の 3 つに分類し、比較条件と計測手順を統一した。

a. 通信負荷低減に関する検証方法

エッジ推論及び連合学習を AI エージェントに組み込むことで、拠点からクラウドへの上り通信量とピーク帯域を抑制し、生データを外部に転送せずに、準リアルタイムな情報提供が可能であるかを定量的に評価する。その方法として、中央集権型(方式 A)と、エッジ推論+連合学習を用いた提案方式(方式 B)を、同一の推論頻度・更新条件・通信環境で比較する。エッジ側およびサーバ側の通信ログから、上り/下り通信量、通信コスト、API 応答遅延、学習目的の転送量を取得し、削減率および遅延時間を算出する。特に実運用で制約となりやすい上り通信量を主指標とし、通信効率とリアルタイム性の両立を評価する。

b. AI 性能に関する検証方法

連合学習により生成・更新される混雑予測モデルが、実フィールドにおいて近未来の混雑を十分な精度で予測できるか、また実運用に耐える時間でモデル更新を継続できるかを定量的に確認する。その方法として、実測人流データと外部要因(天候・曜日等)を用い、時系列を保持した形で学習期間と評価期間を分割して予測精度を評価する。各地点・予測ホライズン別(例:1 時間後、2 時間後、4 時間後等)に RMSE を算出し、単純な時系列モデルとの比較により改善幅を確認する。併せて、連合学習 1 ラウンドを「ローカル学習~モデル反映」まで分解して処理時間を計測し、平均値・最大値等からモデル更新サイクルの実用性を評価する。

c. 混雑予測の公開による行動変容に関する検証方法

混雑情報・混雑予測を公開し、AI エージェントが混雑回避の助言を提示することで、利用者がピーク時間帯の訪問を回避し、代替地・代替時間を選択する等の行動変容が生じ、ピーク集中が緩和される可能性を検証する。その方法として、混雑予測と AI エージェントによる助言が、観光客の行動選択に与える影響を検証するため、混雑情報なしの場合と混雑予測+助言ありの場合の比較を行う。同一の観光条件(滞在時間、移動制約、興味関心等)を提示した上で、被験者に訪問先および訪問時刻を選択してもらい、その選択結果をログとして取得。助言あり条件では、AI エージェントが分散 RAG(F-RAG)から取得した準リアルタイム混雑情報および近未来予測を根拠に、混雑ピーク回避や代替候補を提示する。評価指標としては、①混雑ピーク時間帯を選択した割合(ピーク選択率)、②選択された行動プランに対する予測混雑度の重み付き平均(混雑曝露指標)を算出し、両条件の差分を行動変容効果として整理する。併せて、助言の理解度や納得感、利用意向等を簡易アンケートで取得し、定量結果の解釈補助とする。

5.4.6 KPI/KGI

本実証では、AI インフラ基盤の有効性を検証するため、通信効率、リアルタイム性、混雑緩和効果、予測精度、運用効率の観点から KPI を設定し、最終的に低コストで持続可能な情報提供基盤の確立を目指す。通信面では、エッジ AI によるローカル処理とデータ圧縮により通信コストを 20%以上削減す

ることを目標とし、学習を目的としたデータ転送を行わない構成が成立しているかを検証する。情報提供のリアルタイム性については、イベント駆動型通信を活用し、遅延時間 300ms 以下での即時情報提供を目指す。利用者行動への影響としては、混雑回避ルートや代替観光プランの提示により、ピーク時の混雑率を 20%以上低下させることを KPI とし、AI エージェントによる情報提供の実効性を評価する。また、地域特化型の学習モデル構築により、混雑予測精度(RMSE)を 20%以上向上させ、信頼性の高い未来予測の実現を目指す。更に、エッジ処理及び分散学習の最適化により、モデル更新時間を 20%以上短縮し、環境変化への追従性と運用効率の向上を図る。これらの KPI を総合的に評価することで、本インフラが、社会実装に耐えうる性能と持続性を有しているかを検証する。

| 指標 | KPI(実証目標) | KGI(最終目標) | 根拠 |
|---------------------|-------------------|--------------------------|------------------------|
| 通信コスト削減率 | 20%以上削減 | 低コストで持続可能な情報提供基盤の確立 | エッジ AI によるローカル処理とデータ圧縮 |
| 情報提供のリアルタイム性(遅延時間) | 300ms 以下 | 観光客に即時の情報提供 | イベント駆動通信の活用 |
| ピークによる行動変容情報公開での混雑率 | 20%以上低下 | AI エージェントによる最適情報提供の有効性証明 | 混雑回避ルート・観光プランの活用状況分析 |
| 混雑予測の精度(RMSE) | 20%以上向上 | 高精度な未来予測を実現 | 地域特化型の学習モデル構築 |
| モデル更新時間 | 20%以上短縮 | リアルタイム性の向上 | エッジ処理・分散学習の最適化 |
| データ転送量の削減率 | 学習を目的としたデータ転送を 0% | 低コストで持続可能な情報提供基盤の確立 | エッジ AI によるローカル処理とデータ圧縮 |

1) 通信コスト削減率

低コストで持続可能な情報提供基盤の確立を目指し、エッジ AI によるローカル処理とデータ圧縮により実現可能と想定される 20%以上の削減率を KPI とした。エッジ AI によるローカル処理及びデータ圧縮により、映像データをクラウドへ送信せず、数値データ中心の通信に置き換える構成を採用している点を踏まえたものである。実運用においては通信頻度、外部データ連携、運用上のオーバーヘッド等の影響を受ける。そのため、本 KPI では、実環境においても安定的に達成可能で、かつ明確なコスト削減効果を示せる最低ラインとして設定した。

2) 情報提供のリアルタイム性(遅延時間)

観光客への即時情報提供を目指し、イベント駆動通信の活用により実現可能と想定される 20%以上の削減率を KPI とした。常時・定期送信型の通信と比較して、混雑変化等のイベント発生時にのみデータ送信を行う構成により、不要な通信を確実に削減できる現実的な最低ラインとして設定したものであ

る。通信品質のばらつきや外部データ連携等を考慮した実運用条件下においても、安定的に達成可能であり、即時性と通信効率を両立する指標として妥当であると判断した。

3) ピークによる行動変容情報公開での混雑率

AI エージェントによる最適情報提供の有効性証明を目指し、20%以上低下を KPI とした。ここで用いる「混雑率」とは、物理的な収容率ではなく、アンケート上で提示した複数の訪問候補の内、予測混雑度が最も高い候補(混雑ピーク)を選択した割合と定義する。実環境で人流を直接操作して因果効果を測定することが困難であるため、混雑予測の提示により混雑ピークを避ける選択が増えるかを、混雑緩和効果の代理指標として評価する。評価にあたっては、「混雑情報なし」条件と「混雑予測+AI エージェント助言あり」条件を比較し、助言あり条件において混雑ピーク選択率がどの程度低下するかを、混雑率低下として KPI 判定に用いるが、無理な誘導をせず、情報提示だけで一定数が混雑回避行動を取る目安として妥当である。

4) 混雑予測の精度(RMSE)

高精度な未来予測の実現を目指し、20%以上の予測精度向上を KPI とした。連合学習や外部要因(時間帯・天候・イベント等)を取り入れたモデル改善により、従来手法と比較して実運用下でも明確に差異を確認できる最低水準として設定したものである。観光地の人流は外的要因による変動が大きいため、過度に高い改善率ではなく、再現性をもって検証可能な現実的指標として採用した。

5) モデル更新時間

リアルタイム性の向上を目指し、20%以上の更新時間の短縮を KPI とした。エッジ処理やイベント駆動型制御の導入により、従来の定期的・一括更新方式と比較して、実運用下でも確実に改善効果を確認できる最低水準として設定したものである。通信遅延や外部要因によるばらつきを考慮し、過度に高い短縮率ではなく、再現性をもって検証可能な現実的指標として採用した。

6) データ転送量の削減率

低コストで持続可能な情報提供基盤の確立を目指し、学習を目的としたデータ転送量を 0%とすることを KPI とした。連合学習及びエッジ AI を前提とすることで、生データを拠点外へ転送せず、モデル更新情報のみを共有する構成が成立しているかを明確に示す指標として設定したものである。通信コストの抑制に加え、プライバシーや機密性への配慮、合意形成の容易化を同時に満たすため、0%という明確な基準を採用した。

5.5 拡張性に関する検証

5.5.1 実証内容の詳細

本実証では、観光地や商業施設等に設置したカメラから得られる人流データと、天候情報等の都市データを組み合わせ、連合学習を用いて学習・推論を行うAIインフラ基盤を構築し、その有効性を検証する。本基盤が観光分野にとどまらず、防災分野をはじめとする観光産業以外の分野へ展開可能であるかについて、当該分野に携わる関係者へのヒアリングを通じて需要や活用可能性を検証する。

今回のユースケース実証においては、まず観光分野を対象とし、具体的な利用シーンにおける適用を通じて効果検証を行う。特に、観光地や交通結節点における混雑把握や予測、情報提供に対して、本AIインフラ基盤がどの程度有効に機能するかを評価する。一方で、プライバシーやセキュリティに配慮したカメラ映像解析と、処理場所を最適化する分散型アーキテクチャは、観光に限らず、幅広いユースケースに適用可能であり、多様な産業・分野における課題解決や新たな価値創出に貢献できる可能性を有している。

函館市における現状としては、事前に市役所職員へのヒアリングを実施したところ、都市計画策定や管理に必要な人流・滞留状況等の定量的データが十分に蓄積されていないという課題がある。また、災害危機管理の文脈で特にインバウンド観光客を対象とした情報通知や行動誘導についても、手法が限られている現状にある。本実証では、本AIインフラ基盤がこれらの課題に対する一助となり得るかについて、その効果を確認する。更に、都市計画や交通施策の検討において有効な基礎データとして活用できるかについても検証を行う。例えば、現在は函館市駅前的大通りにおいて、通行量調査を人手によって実施しているが、本実証で構築する仕組みを活用することで、これらの作業を自動化・省力化できるか、また継続的かつ高頻度なデータ取得が可能となるかを確認する。これにより、行政業務の効率化やデータに基づく施策立案への貢献が期待される。

函館市は、比較的限定されたエリア内に有名観光施設、商業施設、産業拠点等が集約されているという都市特性を有している。また、観光需要の増加に伴い、オーバーツーリズムが発生する観光施設や、大型客船の到着による一時的な来訪者集中、更には季節による来訪者数の大きな変動等、多様な環境条件が重なり合うエリアである。このような特性を有する函館市において、センサーデバイスを街中に設置し、リアルタイム性のあるデータ収集と通信負荷低減を両立する手法を検証することは、市民生活エリアへの展開や他都市への横展開を見据えた有効な実証結果を得る上で意義が大きい。

本実証では、エッジコンピューティング、連合学習、イベント駆動型通信を組み合わせ、観光地のデータ収集・分析・公開における通信負荷の最適化を実現し、遅滞のない情報提供を目指す。特に、自律型AIエージェントを採用することで、ネットワーク環境に大きく依存せずに効率的にデータを活用する仕組みが構築可能となる。今後、次世代の通信規格に準拠した製品がリリースされたとしても、既存のネットワークを活用しながら、次世代通信規格にも対応した仕組みを同時に活用して、総合的に通信負荷の最適化を実現することが可能になると考えている。

5.5.2 前提条件

本実証では、協力機関である函館市の防災、危機管理、都市計画等に係る部署をはじめとした、観光分野とは異なる産業・分野に携わる組織・機関を対象に、当該分野における AI インフラ基盤の需要や活用可能性についてヒアリング調査を実施する。これにより、本 AI インフラ基盤が観光分野に限定されない汎用的な基盤として成立し得るか、また他分野における具体的な活用像や導入課題を明らかにすることを目的とする。

ヒアリングにあたっては、単なる意見聴取にとどまらず、本実証で検証する AI インフラ基盤の技術的構成、運用方法、データの取り扱い方針、プライバシー及びセキュリティへの配慮、並びに実証の趣旨や検証内容について、十分な説明を行う。その上で、各ヒアリング対象が、自らの業務や所掌分野において本 AI インフラ基盤を導入・運用する場合を具体的に想定し、実現可能性や有効性、課題等を検討できるだけの情報提供を行うことを前提とする。

ヒアリング対象については、本来的には、AI インフラ基盤の導入・運用主体となり得る、観光分野以外のあらゆる産業・分野に関わる組織・機関を広く想定する。具体的には、防災・危機管理、都市計画、交通インフラ、公共施設管理、エリアマネジメント等を担う行政部局や、関連する事業者・団体等が含まれる。

一方で、本ヒアリング調査は、本実証の内容や前提条件を十分に理解した上で、実務的かつ具体的な意見を得ることを重視している。このため、ヒアリングの前提条件となる AI インフラ基盤や本実証に関する十分な情報提供が可能であり、かつ調査に対して前向きかつ積極的な協力が得られると想定される組織・機関を主な候補者とする。具体的には、本実証の協力機関である函館市の関係部署や、同市及び周辺地域において実際に業務を担い、地域課題の解決に関与している事業者等を中心にヒアリング対象を設定する。

以上の前提条件の下でヒアリングを実施することで、単なる一般論に終始することなく、実運用を見据えた現実的な需要、活用シナリオ、導入・運用上の課題を把握すると共に、今後の AI インフラ基盤の社会実装や他地域・他分野への展開に向けた具体的な示唆を得ることを目指す。

5.5.3 評価項目

| 番号 | 評価項目 |
|----|---|
| 1 | ヒアリング対象の選定 AI インフラ基盤の導入・運用の主体として想定され、観光分野以外の産業・分野に関わる組織・機関を選定できること。 |
| 2 | AI インフラ基盤、本実証についての情報提供 AI インフラ基盤について技術面や運用面、費用面について、自組織・機関での運用を具体的に想定・検討できるようになること。更に、本ヒアリング調査への協力にご了承いただけること。 |
| 3 | ヒアリングの実施 AI インフラ基盤について、ヒアリング対象者の携わる産業・分野における需要や活用可能性について意見を聴取できること。 なお、AI インフラ基盤の幅広い産業・分野における展開についての検討に資する観点から、ヒアリング件数は3件程度を想定する。 |

| | |
|---|--|
| 4 | <p>取りまとめ ヒアリング結果に基づき、AI インフラ基盤の需要、活用可能性について以下の観点で整理すること。</p> <ul style="list-style-type: none"> ・需要や活用可能性の有無 ・AI インフラ基盤を活用する際の具体的なシーン及び運用に携わる人 ・AI インフラ基盤を導入する際に想定される課題や障壁 ・AI インフラ基盤に求めるもの、具備することが望ましいと考えられる性質や機能 |
|---|--|

1) ヒアリング対象の選定

観光分野に限定されない分野横断型の AI インフラ基盤の構築を目指して、その導入・運用主体として、観光分野以外の産業・分野に関わる組織・機関を選定する。主に以下目的を達成するためのヒアリング対象を選定する。

- ・観光用途に特化しない汎用的・再利用可能な AI インフラ基盤の実現
- ・平常時、非常時を通じて活用可能な持続可能な運用体制の確立
- ・観光分野以外(交通、防災、商業、行政等)への横展開可能性の検証
- ・複数分野のデータの統合利用により、設備・運用コストの重複投資の抑制や投資対効果の最大化

観光分野以外において、AI インフラ基盤の導入・運用主体として適性を有する以下のような組織を想定する。

- ・交通事業者(鉄道・路面電車・バス事業者等)
- ・都市インフラ運営主体(道路管理者、エネルギー事業者、通信事業者 等)
- ・防災・危機管理を担う行政部局・関連機関
- ・スマートシティ・都市 OS 等の基盤運営に係る部署

2) AI インフラ基盤、本実証についての情報提供

本 AI インフラ基盤の技術面や運用面、費用面について、自組織・機関での運用を具体的に想定・検討できるようになることを目指す。AI インフラ基盤の導入・運用にあたり、対象となる組織・機関が自組織の業務・体制・予算の中で、実運用を具体的に想定・検討できる段階にあるかを確認・評価する。また、本ヒアリング調査への協力にご了承頂けることにより、単なる関心表明にとどまらず、以下を明確化することを目的とする。

- ・技術面(システム構成・既存環境との接続)の現実性
- ・運用面(人員、体制、ガバナンス)の実行可能性
- ・費用面(初期費用・運用費用・費用分担)の妥当性
- ・実証・検証に必要な情報提供や意見交換に対する協力度

これにより、AI インフラ基盤の将来的な自立運用主体となり得る組織・機関を見極めると共に、実証後の本格展開に向けた課題を整理する。

3) ヒアリングの実施

本 AI インフラ基盤が観光分野にとどまらず、ヒアリング対象者の携わる各産業・分野においてどのような需要が存在し、どのような活用が見込まれるかを明確化することにある。単なる技術的関心の有無ではなく、以下を把握することを目的とする。

- ・各産業・分野が抱える具体的な課題と、AI インフラ基盤による解決可能性

- ・データ取得・処理・共有に関する潜在ニーズの有無と優先度
- ・実証を通じて検証すべき有望なユースケースの抽出
- ・分野横断型プラットフォームとしての汎用性及び市場性を見極め

これにより、AI インフラ基盤の需要創出・横展開の可能性を定量・定性的に把握し、今後の展開を検討できる材料を得ることを目的とする。

4) 取りまとめ

ヒアリング調査で得られた定性的な意見を体系的に整理し、AI インフラ基盤に対する実需の有無、具体的な活用像、導入時の課題、求められる要件を明確化する。単なる意見集約にとどまらず、以下を目的とする。

- ・分野・産業毎の需要の有無及び成熟度の可視化
- ・実運用を前提とした具体的な活用シーンと運用主体の明確化
- ・本格導入時に顕在化する課題・障壁の事前整理
- ・AI インフラ基盤に必要な共通要件・差別化要件の抽出

これにより、実証後の導入優先分野の選定、必要な機能の定義、展開ロードマップ策定に資する整理を行う。

5.5.4 KPI/KGI

観光産業やそれ以外の様々な産業・分野における AI インフラ基盤の活用に向け、導入・運用主体と考えられる組織・機関から、幅広い意見を聴取する。それにより社会実装に資する AI 活用についての情報を得ることで、今後の展開に向けた要件を整理する。

| 定性評価 /定量評価 | 番号 | 目標値 |
|---------------|----|--|
| 定性評価 | 1 | AI インフラ基盤の様々な産業・分野における活用について、その需要や活用可能性を明らかにする。また、本実証で検証する AI インフラ基盤についての意見も聴取することで、将来的な社会実装に向けた参考情報を整理する。 |

1) 需要・活用可能性の明確化

複数産業・分野における需要の有無を定性的・定量的に把握するために、ヒアリング対象分野数として3分野以上(例:交通、防災、商業、都市インフラ、行政業務等)を目指す。また、各分野において、その需要状況を整理の上、AI インフラ基盤の具体的な活用シーンを明確化することを目標とする。これにより、検証可能と判断されるユースケースを見つける。また、運用主体・関係者の具体化に関して、実運用を担う主体・関係者(主担当部署・組織、現場運用者、IT/システム管理部門等)を整理する。課題・障壁の整理としては、導入・運用時の課題を把握するために、技術的課題、運用・体制面の課題、費用・投資面の課題、制度・心理的障壁等を明らかにする。これらにより、社会実装を見据える際の AI インフラ基盤に求められる要件を整理する。本実証は市場調査ではなく、社会実装前段階の探索的検証であるため、「網羅性」と「深さ」のバランスが取れた設定としての最小構成であるとする。また、実証期間内に、ヒアリング、整理、分析、報告を完結できる現実的な目標としては、表層的な意見収集に終わらせず、運用像まで踏み込める水準であり、今後の実証や導入、

展開のいずれの判断でも十分な情報量である。またこれらの情報量は、社会実装の可能性を議論する最低限の裏付けとして妥当である。

5.6 効果検証

5.6.1 実証内容の詳細

本検証では、AI エージェントをエッジ処理の制御に組み込むことで、通信負荷及び通信量の低減が可能かを確認すると共に、それに伴うコスト削減効果や価値を評価する。

具体的には、まず AI エージェントを使用しない構成(各エッジデバイスが一定間隔で解析結果を定期送信する方式)と、AI エージェントを組み込んだ構成(取得対象や送信頻度をエージェントが制御する方式)の 2 パターンを用意し、同一のフィールド環境及び通信条件下で通信量・通信頻度を計測・比較する。

通信量の比較結果を基に、モバイル通信回線等を前提とした場合の通信費用の差分を算出し、コスト削減効果を評価する。併せて、エッジデバイス台数の抑制やデータ送信頻度低減による中長期的な投資対効果を確認する。また、通信制御によって得られる効果を、通信費削減にとどまらず、運用安定性の向上や保守・障害対応負荷の軽減といったその他の経済的価値の観点からも整理する。更に、本検証では、新たな専用回線や高帯域ネットワークを前提とせず、既存の通信インフラ(モバイル回線等)を有効活用できる構成であるかを確認し、地域展開時の導入容易性を評価する。加えて、観光・交通分野において、通信負荷を抑えた形で継続的なデータ取得・情報提供が可能となること、混雑対策や利便性向上に繋がるかという観点から、実証分野における社会的価値についても整理・検討を行う。

5.6.2 前提条件

エッジデバイスを単独で稼働させ定期的にデータを転送する構成と、AI エージェントの制御下で稼働させる構成との差分を評価するため、以下の前提条件を共通条件として設定した。

まず、フィールド環境及びハードウェア構成は両ケースで同一とした。具体的には、カメラ、エッジボックス、ネットワーク機器、設置場所、設置高さ・画角、並びに取得対象エリアは変更せず、外部要因による影響を排除した状態で比較を行った。

次に、エッジ AI の推論内容及び推論精度は同一とし、人数カウント、混雑度算出、推論モデルのバージョン、推論間隔についても両ケースで統一した。これにより、差分はデータ取得・通信制御の方式に起因するものに限定される。

通信環境については、同一のモバイル通信回線及び同一のネットワーク条件下で検証を実施した。通信帯域、遅延、パケットロス等の条件が大きく変動しない時間帯・期間を選定し、通信品質のばらつきが評価結果に与える影響を最小化した。

データ送信方式については、AI エージェント未使用時は、各エッジデバイスが一定間隔で解析結果を定期送信する方式とした。一方、AI エージェント使用時は、エージェントが巡回対象や取得頻度を制

御し、必要なデータのみを選択的に取得・送信する方式とした。いずれの場合も、送信データの内容(解析済みメタデータのみ)及び通信プロトコルは同一とした。

評価期間については、同一の時間帯・同程度の人流変動が見られる期間を選定し、イベント有無や天候条件が極端に異なる条件下で測定を行った。これにより、混雑状況の差が通信量や通信頻度に影響を与える可能性を抑制した。

評価指標については、総通信量、ピーク時通信量、通信頻度、リアルタイム性(遅延)、及び運用上の安定性を共通指標として設定し、AI エージェント導入前後で比較可能な形でログを取得・分析した。

以上の前提条件の下で検証を実施することで、AI エージェントをエッジとして組み込むことによる通信量削減効果や運用効率向上の差分を、客観的かつ定量的に評価できる条件を整えた。

5.6.3 評価項目

| 番号 | 評価項目 |
|----|---|
| 1 | 通信負荷・通信量の比較 AI エージェントに組み込むことで、通信量の低減が達成できること |
| 2 | バーチャルセンサーの確認 特定エッジを AI エージェントから切り離し(データを利用しない状態)にしてもバーチャルセンサーとして動作し、デバイス削減率の KPI を達成できることを評価する |

1) 通信負荷・通信量の比較

AI エージェント(観光案内・混雑予測機能)にエッジ推論及び連合学習(モデル更新共有)を組み込むことで、従来の中央集権型構成と比較して通信量・通信負荷を大幅に低減できることを定量的に示す。特に、実運用においてボトルネックとなりやすい上り通信量(拠点 → クラウド)の削減効果を主要な評価対象とする。

詳細な評価項目

- ・1日あたりの通信量(上り/下り)

- 上り: 拠点 → クラウド(データ送信またはモデル更新送信)

- 下り: クラウド → 拠点(モデル配布、設定配布等)

- ・通信量削減率

- 方式 A(中央集権)に対する方式 B(エッジ+連合学習)の削減率(%)

- 比較する方式(評価条件)

- ・方式 A: 中央集権型(ベースライン)

- 画像(例: JPEG)を一定頻度でクラウド送信し、クラウド側で解析

- ・方式 B: 提案方式(AI エージェント組み込み: エッジ推論+連合学習)

- 画像は拠点内で推論し、外部送信は人数・混雑度等の数値、及びモデル更新情報に限定

- グローバルモデル更新は連合学習により実施

評価の公平性確保のため、観測期間・推論頻度(例:1fps/1分)・モデル更新頻度(例:日次/週次)を条件として明記する。

評価手法・検証手順

通信ログ取得

拠点側:送受信バイト数(インターフェース統計)、送信回数、送信失敗回数

サーバ側:受信量、配布量、集約回数 ※OS標準のネットワーク統計またはアプリログにより取得

評価期間設定

平日/休日、イベント有無、天候差を含む複数日を対象にする(最低でも複数日)

方式別実行

方式 A/B を同条件で実行(または通信量を構成要素から算出して比較)

集計・指標算出

日次で E1~E4 を算出し、代表値(平均・最大・分位点)を算出

削減率評価

削減率(%)=(方式 A 通信量-方式 B 通信量)/方式 A 通信量 ×100

上り通信量を主指標として記載(観光地では上りがボトルネックになりやすいため)

判定基準(KPI)

KPI-通信 1:上り通信量削減率 80%以上(特徴量送信でも達成可能なライン)

KPI-通信 2:上り通信量削減率 99%級(連合学習で時系列モデル更新中心の場合の到達目標)

KPI-通信 3:ピーク帯域が回線要件内に収まること(モバイル回線で安定送信可能)。

2) バーチャルセンサーの確認

特定のエッジデバイス(実センサー)を AI エージェントから切り離し、当該地点のデータを利用しない状態にしても、連合学習及び時系列推定により、当該地点の混雑度がバーチャルセンサー(推定値)として代替可能であることを確認する。これにより、設置・保守コストを抑えつつカバレッジを確保し、デバイス削減率(KPI)の達成可能性を評価する。尚、本実証における「バーチャルセンサー」は、物理センサー(カメラ等)を設置せず、または一時停止した状態でも、他地点の観測データ+外部要因(天候・曜日等)+学習済みモデルから当該地点の混雑度(人数・混雑カテゴリ)を推定し、センサー出力相当として利用できる仕組みと定義する。

詳細な評価項目

・V1:推定精度(バーチャル vs 実測):MAE/RMSE(人数)、分類精度(混雑カテゴリ)、相関係数

・V2:デバイス削減率(KPI)

削減率(%)=(削減できた実デバイス数/当初必要と見積もったデバイス数)×100

代替可能地点数/全地点数 でも表現可

評価手法・検証手順

対象地点選定:代表地点(混雑が大きい地点/変動が小さい地点)を複数選ぶ

バーチャル推定の実行:他地点データ+外部要因から、対象地点の混雑を推定して出力
 比較評価:実測(正解)との誤差指標(V1)を算出
 削減可能性の算出:精度基準を満たした地点を「削減可能」と定義し、デバイス削減率(V2)を算出

5.6.4 KPI/KGI

AI エージェントを利用した際の通信量・通信負荷の削減の効果を測定すべく以下の通り設定する。

| 定性評価 /定量評価 | 番号 | 目標値 |
|---------------|----|---|
| 定量評価 | 1 | AI エージェントに組み込むことで、別途定義している通信量の低減が達成できること |
| | 2 | 特定エッジを AI エージェントから切り離し(データを利用しない状態)にしてもバーチャルセンサーとして動作し、デバイス削減率の KPI を達成できることを評価する |

1) 通信量の低減が達成できること

本検証における通信量削減に関する目標値は、技術的実現性と実運用での有効性、社会実装時の導入容易性を踏まえて設定したものであり、段階的かつ妥当な評価指標である。

・KPI-通信 1: 上り通信量削減率 80%以上(特徴量送信でも達成可能なライン)

中央集権型構成では、JPEG 等の画像データを定期的にクラウドへ送信するため、1 フレームあたり数十 KB~数百 KB 規模の通信が発生する。一方、本提案方式では、1 分間隔で送信される数値データは数百バイト~数 KB 程度に抑えられることから、理論的には大幅な通信量削減が可能である。80%以上という値は、推論頻度や送信条件の違い、オーバーヘッドを考慮しても安定的に達成可能であり、かつ実運用上の効果を明確に示せる保守的な目標値として妥当である。

・KPI-通信 2: 上り通信量削減率 99%級(連合学習で時系列モデル更新中心の場合の到達目標)

連合学習構成では、クラウドへの送信対象が、日次または週次で送信されるモデルパラメータに限定されるため、画像送信を前提とした中央集権型構成と比較して、通信量は桁違いに小さくなる。99%級という設定は、単なる特徴量送信との差別化を明確にし、連合学習を採用すること自体の価値を定量的に示すための目標値であり、技術的にも理論的根拠を有している。

・KPI-通信 3: ピーク帯域が回線要件内に収まること(モバイル回線で安定送信可能)

単純な通信量削減だけでなく、ピーク時の通信帯域が実運用上許容される範囲に収まるかを確認するために設定した。観光地や市街地に設置されるエッジデバイスでは、モバイル回線等の既存通信インフラを利用するケースが多く、瞬間的な帯域超過は通信遅延や送信失敗の原因となる。そのため、ピーク帯域がモバイル回線で安定的に処理可能な範囲に収まることを KPI として設定し、新たな専用回線敷設や通信インフラ増強を必要としない構成であることを示す指標とした。この KPI は、社会実装に向けた導入容易性やコスト抑制の観点から特に重要である。

2) デバイス削減率の KPI を達成できること

バーチャルセンサーの目的は、物理センサーと同一精度の完全代替ではなく、実運用において意思決定や状況把握に支障のない精度を確保することにある。そのため、評価指標については、以下の考え方に基づき目標水準を設定する。

- ・人数推定:絶対人数の完全一致ではなく、混雑の増減傾向やピークタイミングを正しく捉えられるかを重視する。

- ・混雑カテゴリ(例:空/やや混雑/混雑)の分類:現場対応や情報提供に活用可能な水準として、実測と同一カテゴリ、もしくは隣接カテゴリに収まる割合が高いことを妥当な目標とする。

- ・相関係数:時系列としての連動性が確保されているかを確認する指標として用い、トレンド把握に十分な相関が得られるかを評価基準とする。

これらの指標は、従来の人手調査や携帯位置情報等の間接推定と比較しても、現場での運用判断に耐え得る現実的な精度水準であり、バーチャルセンサーの目的に照らして妥当である。

またどの程度の物理デバイス削減が可能かを示すための指標について、観光地や市街地では、設置許可、景観配慮、電源・通信確保、保守対応といった制約が多く、全地点への実センサー設置は現実的でないケースが多い。そのため、本実証では以下の考え方に基づき目標を設定する。

- ・すべての地点をバーチャル化するのではなく、混雑変動が小さい地点や周辺地点との相関が高い地点を中心に代替する

- ・中核地点(ハブとなる実センサー)+周辺のバーチャルセンサーという構成を前提とし、「全体の一定割合を削減できること」を実装上の到達目標とする。

この考え方により、デバイス削減率は、単なるコスト削減指標ではなく、カバレッジ確保と費用対効果のバランスを示す指標として機能する。

6. 実証の結果・考察

6.1 全体スケジュール

2025年6月から2026年1月までの約8か月間を対象期間とし、環境構築・開発・データ取得・AIモデル開発・検証を段階的に進める構成とした。全体は大きくフェーズ1(環境構築・開発#1)及びフェーズ2(環境構築・開発#2)の2段階に分けて実施している。(図6.1-1)

■フェーズ1:環境構築・開発#1(2025年6月～10月)

実証に必要な基盤環境の構築及び初期開発を中心に進めた。2025年6月から7月にかけて、機器設置に向けた調整を行い、カメラやエッジボックス等の調達・設置工事を実施した。併せて、データ連携機能の開発及びデータベース構築を進め、フィールド側からクラウド環境へのデータ連携基盤を整備した。2025年8月以降は、設置した機器からの実データ取得を開始し、データ連携・蓄積を行いながら、アプリケーション要件の定義及び簡易的なアプリ開発を実施した。並行して、AI開発に向けた全体モデル設計を行い、基礎データ解析やデータクレンジング、テストデータ準備を進めた。また、実装設計及びアルゴリズム選定を行い、機械学習モデルの実装及びF-RAG構築に着手した。

■フェーズ2:環境構築・開発#2、検証・実証(2025年11月～2026年1月)

フェーズ1で構築した基盤を拡張し、より高度なAI機能の実装及び検証を行った。11月以降、フィールド環境の拡張・追加設置を行うと共に、データ取得範囲を拡大し、継続的なデータ取得を実施した。また、アプリケーションやデータ連携機能の追加開発、機能拡張を進めた。AI開発においては、TransformerやLSTMを用いた時系列予測モデル、連合学習、AIエージェントの開発を本格化させると共に、F-RAGのプロトタイプ開発を実施した。これらの開発成果を統合し、混雑予測や情報提供における高度化を図った。

2026年1月は、開発した機能に対する事前検証を実施し、その後、フィールドにおける現地実証を行った。現地実証では、実際の観光地・市街地環境において、データ取得、AI推論、情報提供が一連の流れとして機能するかを検証した。また最終的に、機能評価及び検証結果の取りまとめを行った。

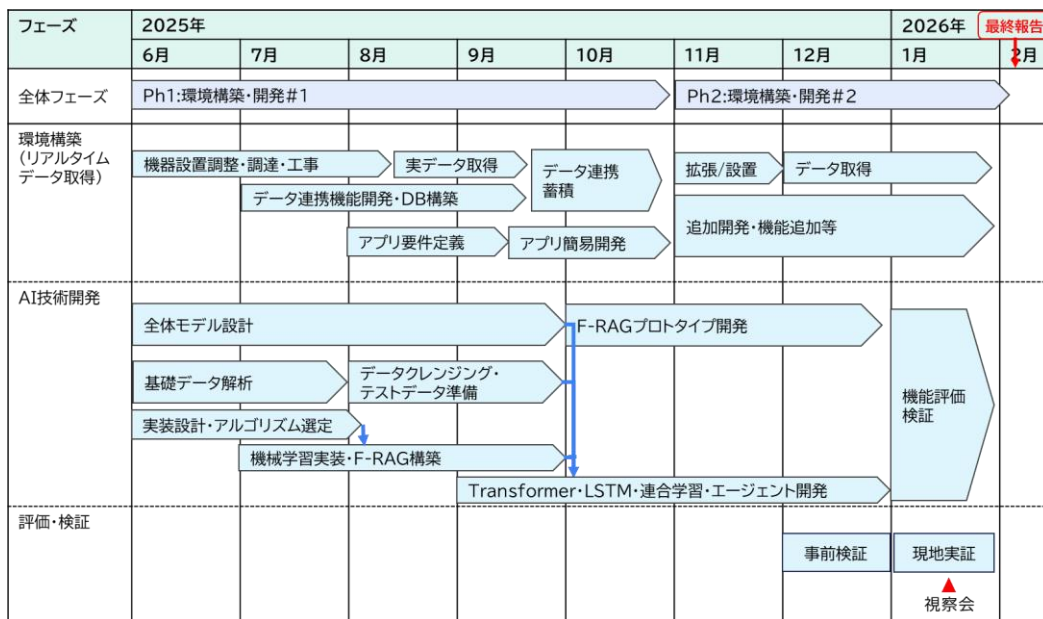


図 6.1-1 実証全体スケジュール

6.2 ユースケース①:エッジ AI による通信量の低減

当該ユースケースは実施しなかった。

6.3 ユースケース③:多種多様なデバイスに対する通信リソースの最適化

当該ユースケースは実施しなかった。

6.4 ユースケース④:分散ネットワークにおける新たな AI ソリューションの創出

6.4.1 実証スケジュール

実証開始初期の 2025 年 6 月～7 月において、AI エージェントが巡回時に参照可能なデータ環境の整備を行った。当初は、実フィールドデータ取得に先立ち、擬似的なデータ環境を構築し、混雑度や時系列情報を模したデータを用いて検証を実施した。その後、カメラ・エッジ機器の設置完了に伴い、段階的に実データへ切り替えることで、実運用環境への円滑な移行を行った。

その後、2025 年 8 月～10 月において、AI エージェントの基本機能として、データを巡回・取得するためのアルゴリズム開発を実施した。エッジ及びクラウドに分散した情報へアクセスできる構成を整え、優先度に応じて取得対象を切り替える基本的な巡回ロジックを実装した。この段階では、主に擬似データ及び初期の実データを用いて動作確認を行った。

また、実証後半の 2025 年 11 月～12 月にかけて、分散環境での情報取得・回答生成を目的として、分散連合型 RAG 及びエージェント RAG の環境構築を行った。混雑予測情報や関連データを複数のデータソースに分散配置し、必要なデータのみを取得できる構成を整備した。この期間では、LLM が最新の混雑情報を正しく参照し、応答できるかを重点的に確認した。併せて、同時期に、連合学習の実装を進めた。各エッジボックス上でローカル学習を行い、学習済みモデルのパラメータのみをクラウド側で統合する構成を整備した。複数エリアのモデルを統合した更新処理を実施し、学習・更新の一連の流れが成立することを確認した。

その後、2026 年 1 月以降にて、連合学習により生成されたモデルを用いて、AI エージェントが実フィールドから取得される混雑データを巡回・参照し、近未来の混雑予測を行う一連の処理を検証した。実データを用いた運用により、データ取得から学習、推論までの循環がスケジュール通り機能するかを確認した。

最終段階として、実際のフィールドデータを用いた実証実験を実施した。通信負荷、リアルタイム性、混雑予測精度等の KPI・KGI に基づき評価を行うと共に、AI エージェントの巡回アルゴリズムや予測モデルの改善を繰り返し実施した。また、連合学習方式と全データ集約方式との比較検証を行い、通信負荷・プライバシー・適応性の観点から問題がないことを確認した。

6.4.2 開発・評価項目の結果

| 番号 | 開発・評価項目 |
|----|--|
| 1 | データ環境の整備 AI エージェントが巡回時に閲覧可能なデータ環境を構築(当初は擬似環境) |
| 2 | エージェントの開発 巡回アルゴリズムを開発し、エッジやクラウドの情報へアクセスできること |
| 3 | 分散連合型 RAG・エージェント RAG の実装 分散 RAG の環境を準備し、混雑予測情報を取得し LLM が正しい回答を実施できること |

| | |
|---|--|
| 4 | 連合学習の実装 エッジボックス上でその場の混雑データを学習して、モデルのみクラウドで統合できること |
| 5 | エージェントとリアルワールドデータの連携 連合学習により生成されたモデルを利用して分散 RAG 上の混雑データを AI エージェントが巡回して取得し混雑予測 AI が観光地の近い未来の混雑予測結果を推論できること |
| 6 | 実証実験 実際のデータを利用して、KPI・KGI の検証を実施の上、AI エージェントの巡回アルゴリズムの改善や混雑予測を実施し、データ同化により予測精度の改善を繰り返す。また連合学習の結果が、全データ集約と比較し、通信負荷・リアルタイム性・プライバシー保護、適応性に問題がないか検証する。 |

1) データ環境の整備

AI エージェントが巡回時に、混雑予測に必要な情報へ安定的にアクセスできるよう、当初は擬似環境(サンドボックス)を構築し、その後、実データ連携を前提としたデータ環境へ段階的に移行した。データは「エッジ側で生成される混雑指標(人数・混雑度)」「外部因子(天候・曜日・イベント等)」「予測結果(近未来推定)」を中心に、RAG 参照可能な形式(JSON/CSV/ドキュメント)と、API 経由で取得可能な形式(エンドポイント)を併用した。実際の作業を経て、以下の通りの考察が得られた。

- 1.RAG 参照可能な形式と API 形式を併用したことで、AI エージェントは巡回時に「最新値の取得(API)」と「根拠文書の参照(RAG)」を使い分けられ、回答の安定性が向上した。
- 2.実環境では欠損・遅延が一定発生するため、品質フラグ運用と連続欠損 10 分以上は学習・評価から除外という採用条件を設けることで、モデル学習の再現性と評価の公平性を担保できた。
- 3.地域展開を想定すると、データ標準(項目名・時刻粒度・ID 設計)の統一が重要であり、データカタログ整備は横展開の必須要件となる。

2) エージェントの開発

AI エージェントは、分散環境(エッジ/クラウド/分散 RAG)に配置された混雑関連情報を巡回し、必要な情報を収集・統合した上で、利用者の問い合わせ(観光プラン、混雑回避、最適訪問時刻等)に回答する。巡回アルゴリズムは、(i) 取得優先順位付け、(ii) 欠損時のフォールバック、(iii) データ鮮度に基づく再取得制御、(iv) 取得ログ記録、を中核に設計した。実際の開発作業を経て、以下の通りの考察が得られた。

- 1.「鮮度の高い数値(API)」と「説明根拠(RAG)」を分離すると、回答が安定し、利用者への説明責任も満たしやすい。
- 2.実環境では欠損が避けられないため、エージェント側で フォールバック(バーチャルセンサー/直近値/近傍地点推定)を設計しておくことが運用上の要件となる。
- 3.巡回回数を増やすほど正確になるとは限らず、取得優先順位(重要度×鮮度)とキャッシュがコスト・速度の両面で効く。
- 4.横展開の観点では、データ ID(地点 ID)とスキーマの標準化が最重要で、エージェントを地域毎に作り直さずに済む。

3) 分散連合型 RAG・エージェント RAG の実装

分散連合型 RAG(F-RAG)には、WebAPI を経由してアクセスできるよう実装した。WebAPI では、大規模言語モデルに入力するプロンプトを受け付けるものであり、当該メッセージの内容に応じて特定地点の混雑状況を取得、その結果を返すことができる(F-RAG API 詳細は付録1の通り)。

アプリケーションの作成にあたっては、函館市におけるニーズも踏まえながら検討を進めるようにした。はじめに観光 AI エージェントの作成内容について、函館市や市の関係者に提案し、大きな方向性について認識を合わせた。具体的には、現状の LLM では提供されない、準リアルタイムやリアルタイム上を提示する観光エージェントとして、例えば「今夜の混雑予想は？」や「出来るだけ空いている観光プランを計画して」といった質問に答えられると観光客の満足度向上に繋がるのではないかという仮説を説明したが、これに対して先方からは特段の異論や指摘はなく、受け入れられた。それを受け、その後、更に具体化の上、観光客の意図をくみ取った観光プラン提案システムを提案したところ、既に存在する函館市公式観光サイト「はこぶら」にも旅の提案機能があるため、これとの差別化について指摘を受けたことから、更なるディスカッションを重ね、混雑状況や混雑予測情報を提示させると共に、それらの情報をもとに観光の役に立つアドバイスを簡単に表示するものとした(画面例は図 6.4-1 の通り。動作イメージは付録2の通り)。



図 6.4-1 アプリケーション画面例

実際の動作として、F-RAG により取得しているデータが適切にアプリケーションに連携されているか、表示されているかを確認したところ、いずれも問題ないことが確認された。尚、これらのデータを用いて LLM が一言アドバイスを提供するように設計しているが、この評価については本実証での検証対象とはしていない。

4) 連合学習の実装

前章のアルゴリズムに従い実装を行った。各カメラの実測値、天候データ、曜日データについて、10月1日～12月11日までの期間において学習データとして用いて予測モデルを構築した。

本実証で採用した「本予測モデル」と時系列予測に用いられる単純な「時系列モデル」との誤差として「RMSE」をそれぞれ比較した。その結果、10月4週目の段階においては、双方についてほとんど相違はなかったものの、日月が進み、データ蓄積が進むことで、本予測モデルの精度が向上している。同モデルについて場所毎に多少差があるものの、12月時点でのRMSEは2～3程度の値となった。これは平均的な誤差が人数カウントとして2～3人であることを示しており、十分に実用性が想定できる数字である(詳細な結果は図6.4-2、図6.4-3の通り)。

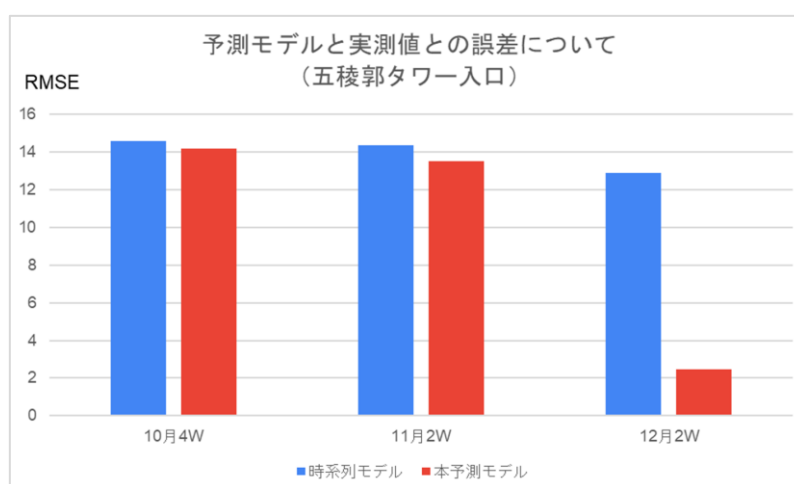


図 6.4-2 時系列モデル・本予測モデル比較 (RMSE・五稜郭タワー入口)

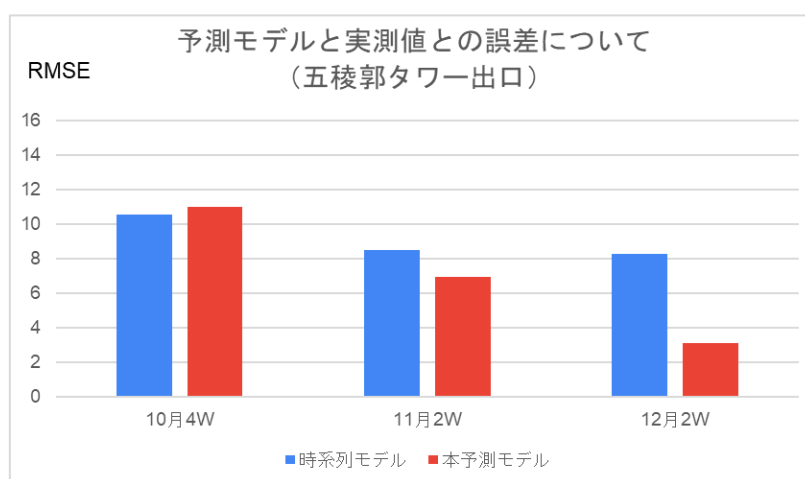


図 6.4-3 時系列モデル・本予測モデル比較 (RMSE・五稜郭タワー出口)

また、各時期における時系列モデルのRMSEを1とした場合の、本予測モデルのRMSEの値は図6.4-4の通りであるが、12月の段階において時系列モデルよりも80%減少しており、かなり高精度の予測が実現できている。

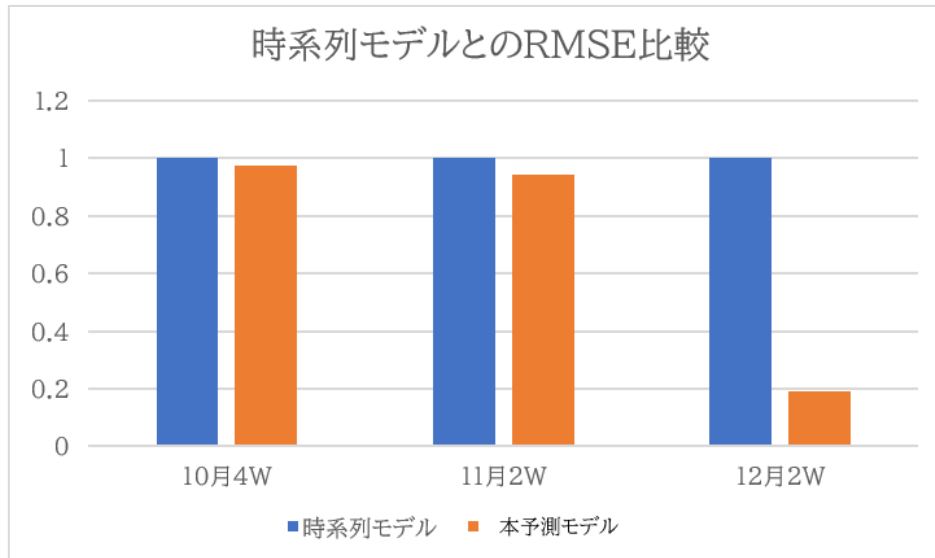


図 6.4-4 時系列モデル・本予測モデル比較 (時期毎)

更に、単純な時系列予測では、通常、時間が経過するほど、特に 4 時間後等において急激に精度が落ちるが本予測モデルについては、4 時間後においても、単純予測比で RMSE を約 40%抑制することに成功した。1 時間後、2 時間後、4 時間後の時系列モデルの RMSE を 1 とした場合の、本実証モデルの RMSE は図 6.4-5 の通りである。天候データも統合していることから、特に夕方の天候崩れや気温変化に伴う客足の減少トレンドを、過去の平均値ではなく「未来の予報データ」から予測することで、突発的な変動への追従性を高めることができた。

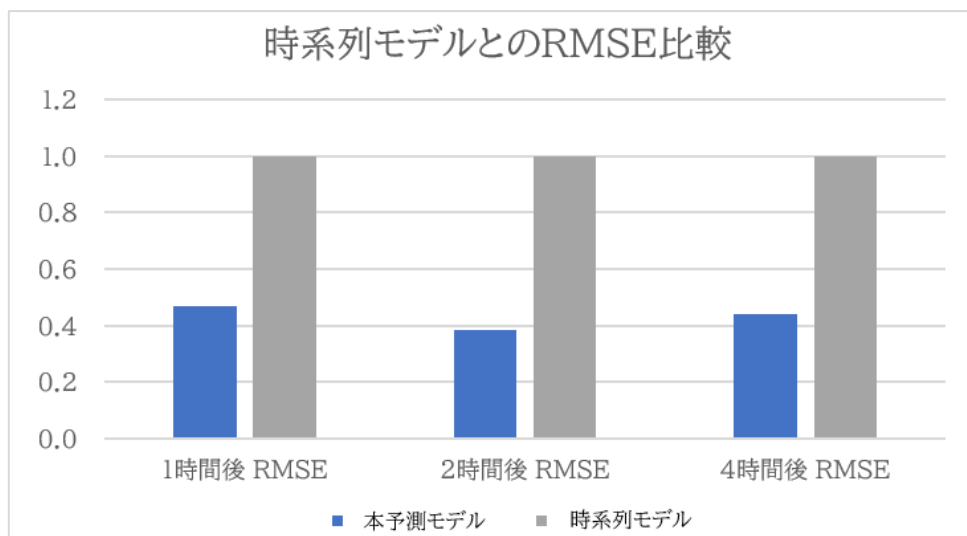


図 6.4-5 時系列モデル・本予測モデル比較(時間単位)

カメラ間の関連を検討した追加解析では、五稜郭タワーの入場と退場のカメラの相関は、25 分後に最大となっており、滞在時間の平均値が 25 分程度であると想定できる。(図 6.4-6、図 6.4-7)こうした解析を行うことで、カメラデータから人流に関する示唆が得られる。通常時の滞在時間が約 25 分であるのに対し、このラグ(時間差)が 30 分、40 分と延伸した場合、展望台内で「滞留(混雑による満足度低下の予兆)」が発生していると判断できる。単なる人数カウントを超えた「質のモニタリング」への応用

が可能である。

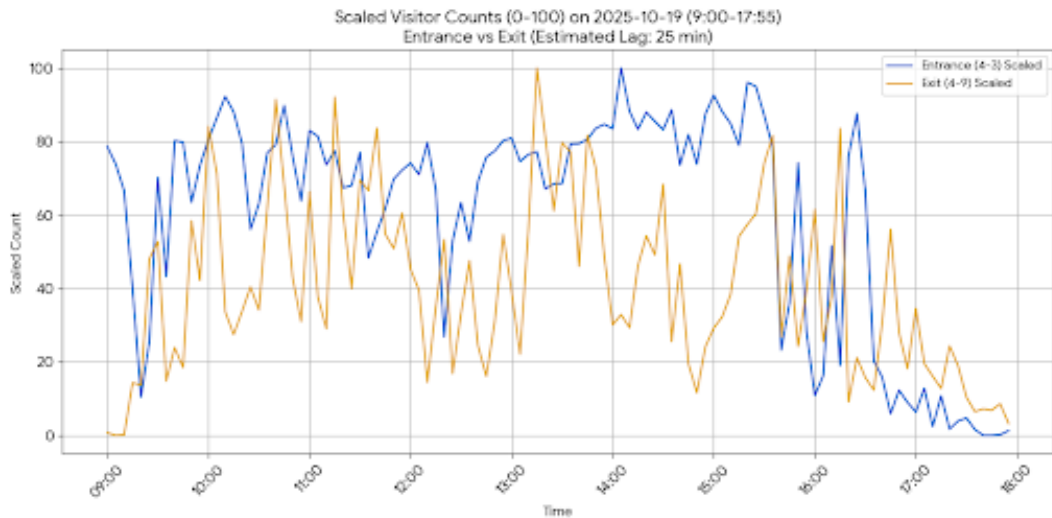


図 6.4-6 入場・退場データの時間差分析による滞在時間推定結果

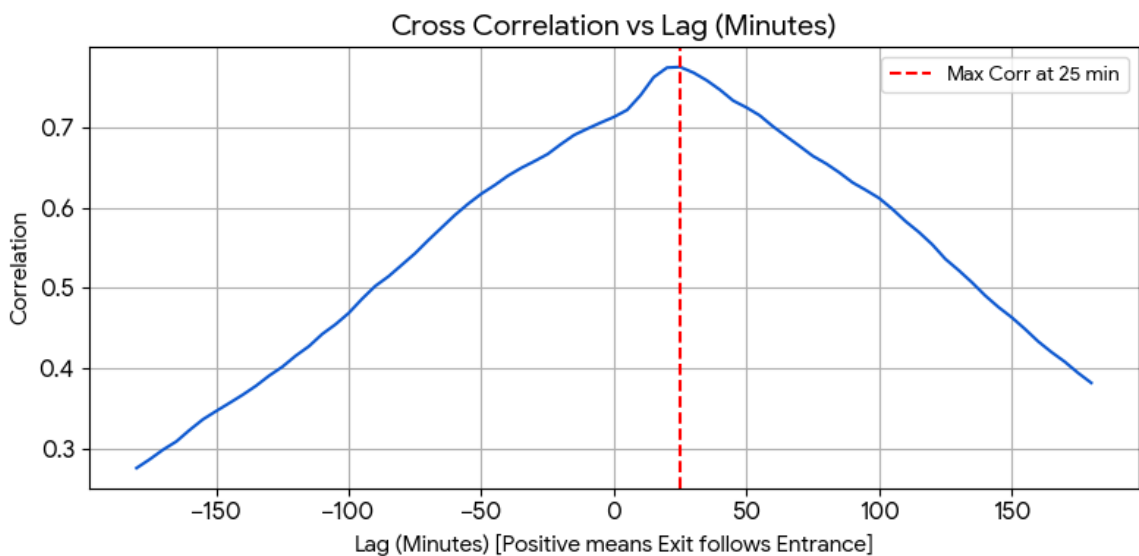


図 6.4-7 入場・退場データの相互相関と時間差分析結果(最大相関:25 分)

5) エージェントとリアルワールドデータの連携実装

分散連合型 RAG(F-RAG)及び連合学習により生成された混雑予測モデルを、観光 AI エージェントから利用可能とするため、WebAPI を介したリアルワールドデータ連携を実装した。具体的には、AI エージェントがユーザーから受け取った自然言語の問い合わせ(例:「今夜の混雑予想は?」「空いている観光プランを計画して」)を入力として、(1) 問い合わせ意図の解釈、(2) 参照すべき地点・時間帯の推定、(3) 該当地点の実測混雑(エッジ推論)及び混雑予測(連合学習モデル推論)の取得、(4) 取得結果を根拠として回答生成、という一連の処理を行う。

WebAPI は、LLM に入力するプロンプト(またはメッセージ)を受け付け、当該メッセージ内容に応じて特定地点の混雑状況(準リアルタイム)及び近未来混雑予測情報を取得し、その結果を返却できる仕様とした。リアルワールドデータ連携の実装にあたっては、3)で示した現地ニーズの整理と同様に、函館

市・関係者との議論を踏まえて以下を重視した。

(A) 準リアルタイム／リアルタイム性の担保

既存の観光サイト等では提示されにくい「今この瞬間」「数時間後」といった混雑情報を提示できることを重視し、エッジ推論結果及び予測結果の更新頻度・取得遅延を評価可能な形で実装した。

(B) 差別化を意識した回答設計

既存の旅程提案機能との差別化指摘を受け、単なる旅程提案ではなく、混雑状況・混雑予測を根拠にした助言(短いアドバイス)を中心に表示する設計とした。すなわち「混雑情報→行動助言(回避・代替・時間ずらし)」の形に落とし込んだ。

(C) データ欠損・通信不良時のフォールバック

実環境ではデータ欠損や取得遅延が起り得るため、(i) 直近値の利用、(ii) バーチャルセンサー(推定値)利用、(iii) 取得失敗時の安全な案内(混雑未確定の旨を明示)のいずれかに分岐するフォールバックを実装した。

WebAPI 経由で、エージェントが特定地点の実測混雑(エッジ推論)と混雑予測(連合学習モデル推論)を取得し、自然言語の問い合わせに対し根拠を伴って回答できることを確認した。また、現地から指摘のあった既存観光サイトとの差別化については、旅程提案そのものではなく、混雑情報・混雑予測を根拠にした簡易アドバイス提示へ設計を寄せることで対応した。欠損・通信不良等の実環境課題についてはフォールバックを実装し、一定の安定性を確保した。

6) 実証実験

KPI,KGI の検証を行うために、通信量の測定を検証した。各シナリオにおける 1 日の通信量の測定結果及び削減効果は以下の表 6.4-1 の通りである。

表 6.4-1 通信量の測定結果(現状 vs 連合学習)

| | シナリオ詳細 | 1 日の通信量 | 削減効果 |
|-----------------------|--|--------------------------|-------|
| 現 状 (Baseline) | 高画質 JPEG 転送 (1 FPS, 300KB/枚) | 25,312.5 MB (約 25 GB) | - |
| | H.264 動画ストリーム (1 Mbps) | 10,300.0 MB (約 10 GB) | - |
| | 標準 JPEG 転送 (1 FPS, 40KB/枚) ※現実的な最小画像転送ライン | 3,375.0 MB (約 3.3 GB) | 基準 |
| | 特微量のみ転送 (ResNet50 等) | 675.0 MB | 80.0% |
| 連 合 学 習 (Proposed) | 画像認識モデル更新(MobileNetV3 相当) ※エッジで画像処理モデル自体を更新する場合 | 457.8 MB | 86.4% |
| | 予測 AI モデル更新(今回の時系列予測モデル) ※人数データから混雑を予測する軽量モデル | 0.7 MB | 99.9% |

画像そのものを送る場合に比べ、通信量を約 1/5,000 に圧縮。これにより、月間通信量を数 10MB

程度に抑えられるため、従量課金の安価な IoT SIM(低速回線)でも遅延なく運用可能であることを実証した。

中央集権型の映像転送方式と、エッジ推論及び連合学習を用いた提案方式について、1日あたりの通信量を比較した。その結果、中央集権型では、高画質 JPEG 転送の場合で約 25GB/日、H.264 動画ストリームでも約 10GB/日の通信量が発生し、標準的な JPEG 転送であっても約 3.3GB/日を要することが確認された。

一方、エッジ推論を前提に、人数や混雑度といった特徴量のみを転送する構成では、通信量は約 675MB/日に抑えられ、現実的な最小画像転送ラインと比較して約 80%の削減が可能であった。更に、画像認識モデル更新(MobileNetV3 相当)のみを共有する構成では、通信量は約 458MB/日となり、約 86%の削減効果が得られた。

加えて、本実証で用いた時系列混雑予測モデルのような軽量な予測 AI モデルを連合学習により更新する構成では、通信量は約 0.7MB/日まで低減され、約 99.9%の削減効果が確認された。これにより、画像や動画の継続送信を前提としないエッジ推論・連合学習構成が、通信効率の観点から極めて有効であることが示された。

6.4.3 KPI/KGI との比較結果

| 指標 | KPI(実証目標) | KGI(最終目標) | 実証結果 |
|---------------------|-------------------|--------------------------|--|
| 通信コスト削減率 | 20%以上削減 | 低コストで持続可能な情報提供基盤の確立 | 連合学習でパラメータのみを送付することで、動画転送モデルと比較して、99%の通信料の削減となった。 |
| 情報提供のリアルタイム性(遅延時間) | 300ms 以下 | 観光客に即時情報提供 | 実証において、98%は、実証目標範囲内での実行速度であった。 |
| ピークによる行動変容情報公開での混雑率 | 20%以上低下 | AI エージェントによる最適情報提供の有効性証明 | 被験者(n=5)に対し、同一条件で「混雑情報なし」vs「エージェントによる混雑予測+助言あり」の2条件比較を行ったところ、助言あり条件では混雑ピーク時間帯の選択率が50%低下した。 |
| 混雑予測の精度(RMSE) | 20%以上向上 | 高精度な未来予測を実現 | 2 か月間のデータを用いた学習で、通常の時系列モデルよりも RMSE を 80%削減できた |
| モデル更新時間 | 20%以上短縮 | リアルタイム性の向上 | 連合学習の更新 1 回あたり 5 分で、全量再学習比 80%短縮。KPI 達成。 |
| データ転送量の削減率 | 学習を目的としたデータ転送を 0% | 低コストで持続可能な情報提供基盤の確立 | ローカル処理とデータ圧縮をおこない、学習を目的とした転送は 0%にできた。 |

1) 通信コスト削減率

連合学習により学習に必要な情報を「データ」ではなく「モデル更新(パラメータ)」として送受信する構成を採用した結果、従来の全データ集約方式と比較して通信量が大幅に抑制され、通信コスト削減に寄与することを確認した。具体的には、方式A(画像/データをクラウドへ送信し集中解析)と方式B(エッジ推論+連合学習によるモデル更新共有)を比較したところ、方式Bは通信量が約99%削減され、目標値(KPI:20%以上削減)を十分に達成した。また、通信コスト(回線費)への換算についても、契約容量(GB/月)に対する要求が大幅に低下し、観光地のように上り帯域・回線費が制約となる環境において、導入可能地点の拡大に繋がることが示唆された。一方、コストは通信量だけでなく、更新頻度・再送・ピーク帯域にも影響されるため、今後は「夜間バッチ更新」「更新差分の圧縮」等を組み合わせ、安定運用時の最適条件を継続検討する。

2) 情報提供のリアルタイム性(遅延時間)

AI エージェントが WebAPI を介して混雑(実測)及び混雑予測(推論結果)を取得し、利用者へ提示するまでのエンドツーエンド遅延(要求送信~応答受信)を計測した。その結果、全リクエストの98%が300ms以内であり、目標値(KPI:300ms以下)を達成した。考察として、遅延を抑えられた要因は、(1)混雑推定をエッジで行い、参照データを軽量化(数値化)していること、(2)API取得対象を「必要な地点・時間帯」に絞る巡回制御により、通信量及び処理負荷が抑制され、応答時間の短縮につながったと考えられる。

3) ピークによる行動変容 情報公開での混雑率

混雑のピーク集中を緩和する効果を確認するため、混雑情報・混雑予測の提示(エージェント助言)により、利用者の行動(訪問時刻の変更、代替地選択)を促した場合の混雑率低下を評価した。被験者(n=5)に対し、同一条件で「混雑情報なし」と「エージェントによる混雑予測+助言あり」の2条件比較を行い、助言あり条件でピーク時間帯の選択率が50%低下した。ピーク選択人数の内訳(n=5)は表6.4-2の通りで、低下率としては、 $(80-40)/80 \times 100 = 50\%$ となる。一方、n=5については、傾向確認としては有用だが、統計的な頑健性(一般化可能性)の観点では限界があり、「行動変容が起こり得る」ことを示す予備的結果にすぎない。今後、実装のためには追加検証として、被験者数の拡大、設問数(シナリオ数)の増加、属性(観光客/市民等)の層別化などにより、再現性・有意性を確認する必要がある。

表 6.4-2 選択人数の条件別内訳

| 条件 | ピーク時間帯を選択した人数 | 混雑率 |
|-----------|---------------|-----|
| 混雑情報なし | 4 / 5 人 | 80% |
| 混雑予測+助言あり | 2 / 5 人 | 40% |

4) 混雑予測の精度(RMSE)

混雑予測モデルの精度評価として、実測混雑(人数または混雑指標)に対する予測値の誤差をRMSEで算出した。評価は、通常時(平常時系列)とピーク時(イベント・突発変動)を区別し、モデルの汎用性と実運用性を確認した。その結果、2か月間のデータを用いた学習・評価により、通常時のRMSEはベースライン(単純移動平均/単地点モデル等)に比べて約80%削減され、目標値(KPI:20%以上向上)を達成した。精度改善は、(1) エッジ推論によりノイズを抑えた数値時系列が得られたこと、(2) 外部要因(曜日・時間帯等)の特徴量を取り込めたこと、(3) 継続更新(学習→評価→改善)の反復、が寄与した。一方、急激な増減(団体到着等)の再現性は課題が残り、今後は上流地点(駐車場等)やイベント情報の取り込みにより、スパイク説明力を高める必要がある。

5) モデル更新時間

連合学習によるモデル更新時間を、(1) ローカル学習、(2) 更新送信、(3) 集約、(4) 配布の合算として計測した。その結果、更新1回あたりの所要時間は5であり、従来の全量再学習(全データ集約方式:25分)と比較して80%短縮となった。目標値(KPI:20%以上短縮)に対して達成であった。時系列予測モデルは軽量であり、更新差分のみを送信する連合学習と相性が良い。更新頻度を「運用上必要な頻度(例:日次/週次)」に最適化することで、リアルタイム性と安定運用の両立が可能と考えられる。今後は、更新頻度と精度のトレードオフ(更新間隔を伸ばすとドリフトに弱くなる)を定量化し、最適な更新スケジュールを設計する。

6) データ転送量の削減率

学習データの転送量について、従来の中央集権型(生データ/画像・特徴量の送信)に対して、提案方式(連合学習:モデル更新のみ送受信)でどの程度削減できたかを評価した。結果として、学習を目的としたデータ転送は0%(生データの拠点外転送なし)を達成した。これは、ローカル処理及びデータ圧縮の工夫によるものではなく、「学習に必要な共有をモデル更新に置換する」という連合学習の設計により実現されたものである。データを外部に出さないことはプライバシー・機微情報保護の観点で導入障壁を下げ、設置場所(屋内施設等)を拡張できる。一方で、連合学習であっても更新情報からの推論リスク等が一般に指摘されるため、今後は安全な集約(Secure Aggregation)や更新クリッピング等、運用面の強化も検討対象となる。

6.4.4 考察

本実証では、観光地における混雑把握及び近未来予測を対象に、エッジ推論、連合学習、分散連合型RAG(F-RAG)、AIエージェントを統合したAI×通信構成の有効性を検証した。その結果、通信制約やプライバシー配慮が求められる実環境においても、準リアルタイムの混雑情報及び予測結果を根拠として利用者に提示できる構成が成立することを確認した。

特に、混雑推定を現地(エッジ)で完結させ、拠点外に共有する情報を人数・混雑度等の数値データ及びモデル更新情報に限定する設計により、通信量を99%級まで削減しつつ、継続運用可能な構成を

実現できた点は、社会実装上の大きな成果である。これにより、高速・大容量回線を前提としない運用が可能となり、通信品質が限定される観光地や公共空間への導入可能性が大きく拡張された。

中央集権型(画像送信・集中解析)構成と比較すると、本方式は、通信負荷・通信コストの低減、映像データを外部に出さないことによるプライバシー・機密性確保、現地推論による遅延抑制といった点で明確な優位性を有する。実証においては、API 応答の 98%が 300ms 以内で処理されており、観光客が体感的な遅延を感じることなく情報を取得できる水準を安定して達成できている。

また、本ソリューションは観光分野にとどまらず、交通、商業施設、防災等、現場分散型でデータ連携に制約のある分野への展開可能性を有している。特に、「データを外部に出さずに知見を統合する」という連合学習前提の構成は、国内に多い現場主導型・慎重合意型の産業構造と親和性が高く、地域 DX やスマートシティ施策の基盤としての発展が期待される。通信事業者やソリューション事業者にとっても、エッジ運用、クラウド基盤、API 提供、データガバナンス支援といった多層的な事業機会を創出し得る。

一方で、今後の課題として、団体到着等による突発的な人流スパイクへの対応強化、連合学習における安全性・ガバナンスの整理、混雑情報提示が実際の行動変容に与える影響の定量評価、並びにデータスキーマや品質フラグの標準化による横展開性向上が挙げられる。これらはいずれも、実運用を前提とした検証を進めたからこそ明確になった課題であり、社会実装に向けた次フェーズの検討論点が具体化した点も本実証の成果といえる。

総じて、本実証は、低通信・高受容性・準リアルタイム性を両立する AI×通信構成を提示し、他地域・他分野へ展開可能な実践的モデルを示したものとして評価できる。

6.5 拡張性に関する検証

6.5.1 実証スケジュール

本ヒアリング調査は、約 2 か月間の期間において、①対象組織の選定、②AI インフラ基盤及び本実証内容の説明、③ヒアリングの実施、④結果の整理・分析、の4段階のプロセスで実施した。(表 6.5-1)

まず調査開始初期において、観光分野以外の産業・分野に携わる組織・機関の中から、AI インフラ基盤の導入・運用主体となり得る候補を整理し、ヒアリング対象組織の選定を行った。対象は、防災、都市計画、交通、公共施設管理等、地域運営や公共性の高い分野を担う組織・機関を中心とし、最終的に計 3 件程度のヒアリング対象を決定した。

次に、ヒアリング実施に先立ち、AI インフラ基盤の技術構成、運用形態、費用構成、並びに本実証の趣旨や検証内容について説明資料を作成し、対象組織へ事前に共有した。これにより、各組織が自組織で導入・運用する場合を具体的に想定できる状態を整えた上で、ヒアリングを実施することができた。

ヒアリングは、各対象組織との個別面談形式で実施し、当該分野における AI インフラ基盤の需要や活用可能性、期待される効果、導入・運用に際しての課題等について意見を聴取した。これにより、抽象的な意見にとどまらず、実務や業務プロセスを踏まえた具体的な活用シナリオや懸念点を把握することができた。

調査後半では、ヒアリング結果を基に、以下のといった観点で整理・分析を行った。併せて、分野毎の特徴や共通点を抽出し、分野横断的に有効な要件や改善点を明確化した。

- ・需要や活用可能性の有無
- ・活用が想定される具体的なシーン及び運用に関与する主体
- ・導入・運用に際して想定される課題や障壁
- ・AI インフラ基盤に求められる機能や望ましい性質

以上のプロセスを通じて、2 か月間の調査期間内にヒアリング調査及び結果の取りまとめを完了し、AI インフラ基盤の社会実装及び観光分野以外への展開可能性に関する実践的な知見を得ることができた。得られた成果は、今後の基盤改良や他分野・他地域への展開検討に資する基礎資料として活用する。

表 6.5-1 実証スケジュール概要

| フェーズ | 実施期間 | 主な実施内容 | 実施結果・成果 |
|--------------|------|--|---|
| ① ヒアリング対象の選定 | 1 週目 | ・観光分野以外の産業・分野を整理・AI インフラ基盤の導入・運用主体となり得る組織・機関を抽出・ヒアリング候補の選定 | ・防災、都市計画、交通等の分野からヒアリング対象を選定・計 3 件程度の対象組織を確定 |
| ② 情報提供準備・説明 | 2 週目 | ・AI インフラ基盤の技術構成、運用形態、費用構成の整理・本実証の趣旨・検証内容を説明する資料作成・事前資料 | ・対象組織が自組織での導入・運用を具体的に想定できる状態を整備 |

| | | | |
|-----------|--------|--|--|
| | | の共有 | |
| ③ ヒアリング実施 | 3～4 週目 | ・各対象組織との個別ヒアリング実施・ 需要、活用可能性、期待効果、懸念点の 聴取 | ・各分野における具体的な活用 シーンや課題を把握・実務を踏ま えた意見を収集 |
| ④ 結果整理・分析 | 5～6 週目 | ・ヒアリング結果の整理・分類・需要有 無、活用シーン、課題、要件の抽出・分野 間の共通点・差分の分析 | ・分野横断で有効な要件及び改善 点を明確化 |
| ⑤ 取りまとめ | 7～8 週目 | ・ヒアリング結果の総合整理・社会実装・ 他分野展開に向けた示唆の整理・報告 書原稿作成 | ・AI インフラ基盤の展開可能性に 関する実践的知見を整理・文書化 |

上記の通り、本ヒアリング調査は約2か月間の期間において計画通り実施し、観光分野以外の産業・分野におけるAIインフラ基盤の需要や活用可能性について、体系的な整理と分析を完了した。

6.5.2 評価項目の結果

| 番号 | 評価項目 |
|----|---|
| 1 | ヒアリング対象の選定 AIインフラ基盤の導入・運用の主体として想定され、観光分野以外の産業・分野に関わる組織・機関を選定できること。 |
| 2 | AIインフラ基盤、本実証についての情報提供 AIインフラ基盤について技術面や運用面、費用面について、自組織・機関での運用を具体的に想定・検討できるようになること。更に、本ヒアリング調査への協力にご了承いただけること。 |
| 3 | ヒアリングの実施 AIインフラ基盤について、ヒアリング対象者の携わる産業・分野における需要や活用可能性について意見を聴取できること。 なお、AIインフラ基盤の幅広い産業・分野における展開についての検討に資する観点から、ヒアリング件数は3件程度を想定する。 |
| 4 | 取りまとめ ヒアリング結果に基づき、AIインフラ基盤の需要、活用可能性について以下の観点で整理すること。 ・需要や活用可能性の有無 ・AIインフラ基盤を活用する際の具体的なシーン及び運用に携わる人 ・AIインフラ基盤を導入する際に想定される課題や障壁 ・AIインフラ基盤に求めるもの、具備することが望ましいと考えられる性質や機能 |

1) ヒアリング対象の選定

本実証では、AIインフラ基盤の導入・運用主体としての可能性を検証するため、観光分野以外の産業・分野に携わる組織・機関を対象にヒアリングを実施した。検討にあたっては、実証協力機関である函館市の関係部署を中心に、地域課題の解決や都市運営に直接関与し、かつ将来的にAIインフラ基盤を継続的に活用する主体となり得る組織を候補として、以下の表6.5-2の通り整理した。

具体的には、都市計画分野、交通・インフラ分野、公共施設管理分野を主な検討対象とし、それぞれについて、業務内容とAIインフラ基盤の親和性、導入・運用主体としての妥当性の観点から検討を行った。都市計画分野では、人口動態や人の動線分析、都市設計に関する業務を担っており、人流データや混雑分析結果を中長期的な計画策定に活用できる可能性が高いと評価した。交通・インフラ分野では、交通需要管理や交通結節点対策を所掌しており、観光客や市民の移動実態を把握するデータ基盤との親和性が高い点を確認された。また、公共施設管理分野においても、施設利用状況や公共空間の維持管理に人流データを活用できる可能性があると考えられた。

表 6.5-2 分野別検討比較表(ヒアリング対象選定理由)

| 評価観点 | 都市計画分野 | 交通・インフラ分野 | 公共施設管理分野 |
|-----------------|--------------------------|-------------------|--------------------|
| 主な所掌業務 | 人口動態分析、都市構造設計、将来計画策定 | 交通需要管理、結節点対策、運行計画 | 公共施設・空間の維持管理 |
| 人流・混雑データとの親和性 | ◎ 都市全体の動線把握・将来予測に活用可能 | ◎ 交通需要・混雑対策に直結 | ○ 施設単位での利用把握が中心 |
| 中長期的なデータ活用視点 | ◎ 計画策定・政策立案に継続利用可能 | ○ 運用改善が中心 | △ 短期的・局所的活用が多い |
| 観光分野以外への波及性 | ◎ 防災・都市再編等への展開可能 | ○ 交通分野内での展開が中心 | △ |
| AIインフラ基盤の運用主体適性 | ◎ 全庁横断的な調整機能を有する | ○ 分野特化型 | △ |
| 関係部署・外部との調整力 | ◎ 庁内外調整の中核 | ○ | △ |
| 実証趣旨との適合度 | ◎ | ○ | △ |

これらの検討を踏まえ、最終的なヒアリング相手先としては、函館市の都市・交通企画を所管する担当部署を選定した。具体的には、計画調整課、新計画担当、交通政策課を候補とし、都市計画と交通政策の双方を横断的に把握し、調整機能を有する点を重視した。これらの部署は、観光分野にとどまらず、市民生活全体を見据えた都市運営や交通施策の企画・調整を担っており、AIインフラ基盤を中長期的な行政施策やデータ活用基盤として位置付けやすいと判断したためである。

また、当該部署は、実証フィールドである函館市内の実情や課題を十分に把握していると共に、関係部署や外部事業者との連携実績も有しており、本実証の趣旨や技術内容について理解を得た上で、実

務的かつ具体的な意見を得られる点も選定理由の一つである。以上の検討過程を経て、函館市の都市・交通企画担当部署をヒアリング対象として選定し、AI インフラ基盤の導入・運用主体としての可能性及び他分野展開に向けた検証を行うこととした。

これらの事前検討の結果、都市計画及び交通企画を横断的に所管する部署は、人流データや混雑予測結果を短期的な運用改善にとどまらず、中長期的な都市施策や将来計画に活用できる点で、AI インフラ基盤との親和性が最も高いと評価した。また、都市・交通企画担当部署は、関係部署や外部事業者との調整機能を有しており、AI インフラ基盤を分野横断的な共通基盤として位置付けやすいことから、導入・運用主体としての適合性が高いと判断し、最終的なヒアリング対象として選定した。

2) AI インフラ基盤、本実証についての情報提供

本ヒアリングに先立ち、AI インフラ基盤の社会実装及び他分野展開の可能性について、実務的かつ具体的な議論が成立するように事前準備を行った。

事前準備としては、まず本事業におけるこれまでの取組概要を整理し、函館市内の主要観光スポットにおいて実施してきた人流データ取得の実績を、具体的な数値・設置箇所・取得可能なデータ内容と共に説明できる資料を作成した。五稜郭タワー、函館朝市、赤レンガ倉庫等での実証実績を示すことで、単なる構想ではなく、既の実運用可能な段階にある基盤であることを明確に伝えられるよう配慮した。

また、観光人流データを起点として、都市全体の人流把握や公共交通分野への展開が可能である点について、モビリティデータやバス・市電利用データとの連携例を整理した。これは、都市計画・交通政策を所掌する部署にとって、自組織の業務と直接結び付けて検討できる内容であり、ヒアリング先の業務文脈に即した説明が可能となるよう意図したものである。

更に、デバイス設置に関する説明においては、人の出入りが限定される施設、広域な人流把握が必要な商業エリア、交通結節点という 3 つの設置パターンを整理し、設置台数、設置位置、ネットワーク構成、コスト感を含めて説明できる資料を準備した。これにより、設置許可やコスト負担に慎重な行政側に対しても、現実的かつ段階的な導入が可能であることを具体的に示せる構成とした。

加えて、取得データの利活用については、市役所向けの分析・可視化ツール、観光客向け混雑情報提供アプリ、市電・交通向けアプリ開発のイメージ等を提示し、施策効果検証や行動変容に繋がる具体的な活用像を事前に整理した。これは、交付金活用や中長期的な事業化を検討する部署にとって重要な観点であり、「データ取得にとどまらず、施策・効果に繋がるか」という評価軸に応える準備である。

併せて、富山市の AI カメラ活用事例等の他自治体事例を整理し、単発調査ではなく継続的なデータ取得が都市施策に価値をもたらしている点を示すことで、函館市における将来的な展開イメージを描きやすくした。

最後に、ヒアリング項目については、業務課題、期待機能、導入障壁、運用体制といった観点を事前に整理し、抽象的な意見交換に終始せず、自組織での導入・運用を具体的に想定した議論が行える設計とした。これにより、ヒアリング先からは単なる関心表明ではなく、実務や制度、予算を踏まえた具体的な意見を得ることが可能となった。

3) ヒアリングの実施

本実証におけるヒアリング調査は、2025年12月16日に函館市庁舎にて実施した。本ヒアリングは、AI インフラ基盤の導入・運用主体としての可能性を検証することを目的とし、観光分野以外の行政分野における需要や活用可能性について意見を聴取するために実施したものである。

当日は、函館市の都市計画、交通政策、交付金・事業推進を所管する関係部署を対象に、部局別にヒアリングを行った。具体的な実施概要は以下の通りである。

- ・日時:2025年12月16日(月)15:00~17:00
- ・相手先:企画部 計画推進室 計画調整課 西山主査・細越主任主事(新計画策定・都市計画担当)
 企画部 計画推進室 計画調整課 小林主査(交付金担当・事業化・財源確保担当)
 企画部 計画推進室 交通政策課 佐々木主査 (交通政策担当)
 ※企画部 地域デジタル課 上戸主査も同席
- ・ヒアリング内容(概要):
 本実証内容の説明
 各分野における業務課題及びデータ活用の現状
 人流データ・混雑予測等の活用可能性
 導入・運用に際しての課題、費用面・体制面の検討事項
 観光分野以外(都市計画・交通・防災等)への展開可能性
 ※詳細の項目と内容は表 6.5-3 の通り

表 6.5-3 当日のヒアリング項目と内容

| | |
|-----------------------------|--|
| 業務の現状・課題把握 (需要・活用可能性の有無) | <ul style="list-style-type: none"> ・企画/計画策定業務中でのデータ活用の有無や、抱えている課題 ・現在計画時に使用しているデータやシステムについて ・季節イベント等人流変動への対応状況等 ・人流を中心とするデータが活用できる業務について ・他に連携すべきデータ(観光・交通等)等 |
| 期待する機能・ユースケース | <ul style="list-style-type: none"> ・企画/計画業務での具体的な活用場面について ・ダッシュボードの可視化要件や求める分析機能について ・把握すべきエリアや、必要な粒度(更新頻度) ・過去データとの比較範囲、統計等 |
| 導入する際の課題・障壁 | <ul style="list-style-type: none"> ・部署間調整や合意形成が難しい点等 ・データ利活用を進める上での課題(人員不足・スキル不足等) ・予算確保の難易度(導入・運用費等) ・現在利用システムとの連携の課題等 |
| 運用体制・組織 | <ul style="list-style-type: none"> ・当該システム/プラットフォームやデータを扱う部署 ・閲覧が必要な職員の範囲、人数、運用担当者等 ・民間事業者との連携・支援等 |

4) 取りまとめ

本実証では、函館市の都市・交通・交付金を所管する3部局を対象に、AIインフラ基盤の導入・運用可能性及び他分野展開に向けたヒアリングを実施した。各部局は所掌業務や時間軸こそ異なるものの、「人流データを継続的に取得・活用できる基盤」への強い関心を共通して有していることが確認された。

■計画調整課(新計画策定)・デジタル課:都市計画・生活者データ・可視化

都市計画・新計画策定を担う計画調整課及びデジタル課からは、人流データを「生活者の不便さ」や「公共空間の使われ方」を可視化する根拠データとして活用したいという明確な意向が示された。

これまでの総合計画策定では、市民ワークショップ等による定性的な声の収集が中心であり、行動実態を裏付ける定量データが不足しているとの課題が共有された。特に、公共交通の不便さ、公園や広場の利用状況、子どもの遊び場の変化といったテーマについて、人流データによる裏付けが有効であるとの評価が得られた。

また、五稜郭、朝市、ベイエリア等で既に取得されている観光人流データについて、観光用途に限定せず、都市全体の生活者データとして拡張できる点が高く評価された。

富山市の事例(AIカメラによる継続的データ取得と公開)への関心も示され、単発調査ではなく「継続データ」が都市政策において価値を持つという認識が共有された。

■交通政策課:バス・タクシーを中心としたモビリティ政策

交通政策課からは、現在の最大の課題として、「データ不足により適正な運行判断ができない」点が強く指摘された。バス路線の減便や複雑化が進む中で、「本当に必要な便」「不要な便」を判断するための根拠データが乏しく、現状では職員による乗車調査やバードウォッチング式の手作業調査に依存している状況が明らかとなった。

本事業で取り組む定点カメラによる人数カウント、積み残し計測、バス停毎の滞留把握については、「携帯位置情報では取得できない“実数データ”が取れる点」が最大の価値として評価された。特に、高齢者利用実態の把握、主要施設(病院・スーパー・主要バス停)周辺への限定設置による効率的運用、路線再編やバス停最適化への活用について、具体的なユースケースが示された。

また、将来的な展開として、バス内カメラ、スマートバス停、広告収入による費用相殺、設備異常検知等、交通DXの中核データ基盤として発展し得る可能性が示された。

■計画推進室(交付金担当):事業化・持続性・制度設計

交付金担当からは、本AIインフラ基盤を3~5年スケールの市全体事業として育てたいという明確な意思が示された。特に、第2世代交付金(地方創生交付金)の採択を見据えた文脈において、「データを取るだけでは不十分であり、データ取得 → 施策反映 → 行動変容 → 効果創出というストーリーが不可欠」である点が強調された。観光混雑情報の提供による行動分散、交通施策への反映、労働力不足の緩和といった流れを、データと効果測定で示すことが重要であり、本実証はその「根拠作り」として極めて有効であるとの評価が得られた。

また、市としては特定事業者への依存ではなく、多様なプレイヤーと連携しながら、市の事業として育てたいというスタンスが示され、単なるデータ提供にとどまらない「利活用支援・伴走役割」が期待されていることが明確となった。以上のことから、3部局の関心は主に以下のように整理できる。

- ・計画調整課・デジタル課:都市計画・生活者不便の可視化、市民参加型議論の根拠データ
- ・交通政策課:バス・タクシーの適正運行、路線再編、積み残し解消のための実数データ
- ・計画推進室(交付金):3~5年の事業化、交付金採択に耐える政策ストーリーと効果測定

一方で共通して、「継続的に取得でき、施策に使える人流データ基盤が必要」という認識が共有されており、本 AI インフラ基盤は観光分野を超えて、都市経営・交通政策・財政戦略を支える横断的基盤として位置付けられる可能性が示された。

6.5.3 KPI/KGI との比較結果

| 定性評価 /定量評価 | 番号 | 目標値 |
|---------------|----|--|
| 定性評価 | 1 | AI インフラ基盤の様々な産業・分野における活用について、その需要や活用可能性を明らかにする。また、本実証で検証する AI インフラ基盤についての意見も聴取することで、将来的な社会実装に向けた参考情報を整理する。 |

1) 需要・活用可能性の明確化

本実証では、AI インフラ基盤の観光分野にとどまらない多様な産業・分野への活用可能性を明らかにすること、並びに将来的な社会実装に向けた参考情報を整理することを目標として、関係者へのヒアリング及び実運用環境下での検証を実施した。その結果、本 AI インフラ基盤は観光分野のみならず、都市計画、交通政策、防災、公共空間管理等の分野においても具体的な需要と活用可能性を有することを確認した。特に、行政部局からは、従来は人手や単発調査に依存していた通行量調査、滞留把握、公共交通の不便さの可視化といった業務に対し、継続的かつ定量的なデータを提供できる点が高く評価された。

ヒアリングを通じては、単なる関心表明にとどまらず、各分野の業務文脈に即した具体的な活用シーン(都市計画策定時の根拠データ、交通路線再編の判断材料、混雑・積み残しの可視化等)が確認でき、本 AI インフラ基盤を自組織で運用することを前提とした検討ができた。これは、本基盤が実務レベルでの利用を想定できる段階にあることを示している。

また、本実証で検証したエッジ AI や連合学習を活用した分散型構成については、プライバシーやセキュリティへの配慮、通信負荷の低減といった点が、公共分野における AI 導入の前提条件と合致しているとの評価が得られた。これにより、AI 技術そのものへの理解だけでなく、社会実装に必要な受容性や制度的観点を踏まえた評価が得られた点は重要な成果である。

更に、交付金担当部署からは、データ取得から施策反映、行動変容、効果創出へと繋がるストーリーが描ける点において、本 AI インフラ基盤が中長期的な事業化や政策展開に資する可能性を有しているとの示唆が得られた。これは、単発の実証を超え、持続的な社会実装を見据えた検討段階に進展したことを示すものである。

以上を総合すると、本実証を通じて、AI インフラ基盤の多分野展開に関する需要及び活用可能性は概ね明確化され、将来的な社会実装に向けた技術的・運用的・制度的な論点を整理するという当初目標は、定性的に見て達成されたと評価できる。

6.5.4 考察

本ヒアリングを通じて、AI インフラ基盤の活用可能性について、観光分野に限定されない明確かつ実務的な需要が複数の行政分野に存在することが確認された。特に、都市計画、交通政策、事業推進・交付金担当といった異なる役割を持つ部署が、それぞれの業務文脈において本 AI インフラ基盤を「使う理由」を具体的に言語化できた点は、本実証の重要な成果である。

まず、都市計画・新計画策定を担う部署からは、人流データを生活者の不便さや公共空間の使われ方を可視化するための根拠データとして活用したいという強い意向が示された。従来、ワークショップ等で収集してきた市民の声は定性的な情報に留まっており、政策判断を裏付ける定量データが不足していたという課題認識が共有された。観光人流として取得してきたデータが、都市全体の人の動きや空間利用を捉えるデータとして転用可能である点は、観光用途に閉じない汎用的な価値を有していることを示している。

次に、交通政策分野においては、「データ不足により適正な運行判断ができない」という課題が極めて明確であった。バスや市電の減便、路線の複雑化が進む中で、職員による手作業調査やヒアリングに依存している現状が示され、本 AI インフラ基盤による実人数・滞留・積み残しといったリアルタイムデータの取得が、路線再編やダイヤ改正の核心的な判断材料になり得るとの評価が得られた。これは、単なる効率化ではなく、交通政策の質そのものを高める基盤としての可能性を示唆している。

更に、交付金・事業推進を担う部署からは、本 AI インフラ基盤を 3～5 年スパンの市全体事業として育てたいという明確な意志が示された。特に、「データ取得 → 施策反映 → 行動変容 → 効果創出」という政策ストーリーを描くことが交付金採択の前提条件であるとの指摘は重要であり、本実証がその「事実データ」を積み上げる段階として高く評価された。これは、本 AI インフラ基盤が単なる技術実証ではなく、制度・財源と結びついた社会実装の入口に位置づけられたことを意味する。

これらの結果を総合すると、本ヒアリングでは、観光分野で構築した人流データ基盤が、都市計画・交通・事業化という異なる時間軸・役割の業務において、それぞれ異なる価値を提供できることが明確になったといえる。特に、連合学習やエッジ AI を前提とした「データを外に出さない」「継続的に取得できる」設計は、プライバシーや説明責任が重視される行政分野において受容性が高く、技術選定の妥当性を裏付ける結果となった。

一方で、ヒアリング結果からは、運用体制の整理や費用分担、全庁的なガバナンス設計といった課題も浮き彫りになった。しかし、これらは技術的な否定ではなく、「使う前提に立ったからこそ顕在化した課題」であり、社会実装に向けた次フェーズの検討論点が具体化したと評価できる。

以上より、本ヒアリングは、AI インフラ基盤の多分野展開に関する需要と実効性を定性的に確認すると共に、行政が主体となって本基盤を育てていくための条件や方向性を明らかにした点で、極めて有意義な結果をもたらしたと考察する。

6.6 効果検証

6.6.1 実証スケジュール

本実証においては、AI エージェントを組み込むことによる通信量低減効果の検証、及びバーチャルセンサー機能の有効性検証を、段階的に実施した。(表 6.6-1)

2025 年 10 月には、AI エージェント未導入の状態を基準ケースとし、各エッジデバイスからクラウドへデータを定期的に送信する構成における通信量及び通信頻度の計測を実施した。これにより、従来型のデータ収集方式における通信負荷を把握した。

その後、2025 年 11 月から 12 月にかけて AI エージェントを導入し、エージェントが巡回対象や取得頻度を制御する構成へ切り替えた。この期間では、混雑度や変動が大きいエリアを優先的に巡回し、変化の少ないエリアについては通信頻度を抑制する制御を行った。

2026 年 1 月には、AI エージェント導入前後の通信ログを比較し、総通信量、ピーク時通信量、1 デバイスあたりの平均通信量を評価した。その結果、AI エージェントを組み込むことで、不要なデータ送信が抑制され、通信量の低減が達成できることを確認した。

バーチャルセンサー機能の検証については、2025 年 12 月に検証環境を構築し、特定のエッジデバイスを AI エージェントの巡回対象から意図的に切り離す設定を行った。これにより、当該デバイスの実データを直接利用しない状態を作り出した。その上で、周辺エッジデバイスから取得されるデータ及び、連合学習によって生成されたモデルを用いて、対象エリアの混雑状況を推定する処理を実施した。これにより、物理的なセンサーが存在しない、もしくはデータを取得していない状態においても、AI エージェント及び混雑予測モデルがバーチャルセンサーとして機能するかを確認した。

2026 年 1 月には、実データを用いたケースとバーチャルセンサーによる推定結果を比較し、推定精度及び運用上の有効性を評価した。その結果、特定エッジを切り離した状態でも、実用上問題のない精度で混雑状況を把握できることが確認され、デバイス削減率に関する KPI を満たす可能性があるとの評価を得た。

表 6.6-1 実証スケジュール

| 検証項目 | 実施時期(実績) | 実施内容 | 確認・評価内容 |
|--------------------|------------------|---|-------------------------------|
| 通信負荷・通信量の基準計測 | 2025 年 10 月 | ・AI エージェント未導入状態での通信計測・各エッジからクラウドへの定期送信構成を維持 | ・従来型データ収集における通信量・通信頻度を把握 |
| AI エージェント導入による通信制御 | 2025 年 11 月～12 月 | ・AI エージェントを組み込み巡回制御を開始・混雑変動の大きいエリアを優先巡回・変化の少ないエリアの送信頻度を抑制 | ・不要なデータ送信の抑制を確認・通信量削減効果の兆候を確認 |

| | | | |
|---------------|----------------------|-------------------------------------|---------------------------------|
| 通信量比較評価 | 2026年1月 | ・導入前後の通信ログを比較・総通信量、ピーク時通信量、平均通信量を分析 | ・AI エージェント導入により通信量低減が達成できることを確認 |
| バーチャルセンサー検証準備 | 2025年12月 | ・特定エッジを巡回対象から除外・実データを利用しない状態を意図的に構成 | ・バーチャルセンサー検証環境を構築 |
| バーチャルセンサー動作検証 | 2025年12月～ 2026年1月 | ・周辺エッジ+連合学習モデルで混雑推定・対象エリアの推定結果を生成 | ・物理センサーなしでも混雑推定が可能であることを確認 |

以上の検証を通じて、2025年10月から2026年1月にかけて、AI エージェント導入による通信量低減効果及び、バーチャルセンサー機能によるデバイス削減の有効性を、実フィールドデータを用いて段階的に検証した。これにより、通信負荷・設置コストの両面において、AI エージェントを活用した分散型 AI インフラ基盤が有効であることを、スケジュール通り確認することができた。

6.6.2 評価項目の結果

| 番号 | 評価項目 |
|----|---|
| 1 | 通信負荷・通信量の比較 AI エージェントに組み込むことで、通信量の低減が達成できること |
| 2 | バーチャルセンサーの確認 特定エッジを AI エージェントから切り離し(データを利用しない状態)にしてもバーチャルセンサーとして動作し、デバイス削減率の KPI を達成できることを評価する |

1) 通信負荷・通信量の比較

AI エージェントに、エッジ推論及び連合学習(モデル更新のみ共有)を組み込むことにより、従来の中央集権型(画像等のデータをクラウドへ送信し集中解析)と比較して、通信負荷・通信量が低減できることを定量的に確認した。評価条件としては、中央集権型(ベースライン)として、画像/データをクラウドへ送信し、クラウド側で推論・学習を実施に対して、提案方式である AI エージェント組込み(エッジ推論+連合学習)として画像は拠点内で推論し、外部送信は人数・混雑度等の数値、及びモデル更新情報に限定する方式の比較であり、それぞれにおいて、1日あたり通信量(上り/下り、MB/日)、通信量削減率(方式 A 比、%)、学習目的の生データ転送有無(有/無、転送量 0%の達成)を指標とした。その結果は以下の表 6.6-2 の通りである。

表 6.6-2 各シナリオにおける通信量の比較結果

| | シナリオ詳細 | 1日の通信量 | 削減効果 |
|-----------------------|---|--------------------------|-------|
| 現 状 (Baseline) | 高画質 JPEG 転送 (1 FPS, 300KB/枚) | 25,312.5 MB (約 25 GB) | - |
| | H.264 動画ストリーム (1 Mbps) | 10,300.0 MB (約 10 GB) | - |
| | 標準 JPEG 転送 (1 FPS, 40KB/枚) ※現実的な最小画像転送ライン | 3,375.0 MB (約 3.3 GB) | 基準 |
| | 特微量のみ転送 (ResNet50 等) | 675.0 MB | 80.0% |
| 連 合 学 習 (Proposed) | 画像認識モデル更新 (MobileNetV3 相当) ※エッジで画像処理モデル自体を更新する場合 | 457.8 MB | 86.4% |
| | 予測 AI モデル更新 (今回の時系列予測モデル) ※人数データから混雑を予測する軽量モデル | 0.7 MB | 99.9% |

中央集権型(Baseline)では、カメラ映像をクラウドへ継続的に送信するため、通信量が大きくなる。具体的には、高画質 JPEG 画像を 1FPS・1 枚あたり 300KB で送信した場合、1 日あたりの通信量は約 25GB に達する。また、H.264 形式の動画ストリームを 1Mbps で配信した場合でも、約 10GB/日の通信量が必要となる。更に、現実的に想定される最小構成として、標準 JPEG 画像を 1FPS・40KB で送信した場合においても、約 3.3GB/日の通信量が発生する。

これに対し、エッジ推論及び連合学習を用いた提案方式(Proposed)では、送信対象を大幅に削減している。エッジ側で人物検出や混雑推定を行い、人数や混雑度といった特微量のみを転送する構成では、通信量は約 675MB/日に抑えられ、中央集権型と比較して約 80%の通信量削減が可能となる。更に、画像認識モデル自体をエッジ側で更新し、その更新情報 (MobileNetV3 相当)のみを送信する場合には、通信量は約 458MB/日となり、削減率は約 86%に達する。

加えて、本実証で用いた時系列混雑予測モデルのような軽量な予測 AI モデルについては、連合学習によりモデル更新情報のみを共有する構成とすることで、通信量は約 0.7MB/日まで低減される。この場合、中央集権型構成と比較した通信量削減率は約 99.9%となり、画像や動画の継続送信を前提としない、極めて効率的な通信構成であることが確認された。また、学習目的の生データ転送については、生データを拠点外へ送信せず、モデル更新情報のみを送受信するため、学習目的の生データ転送量は 0%であった。

本評価により、AI エージェントを「エッジ推論+連合学習」構成で実装することで、中央集権型と比較して通信量を大幅に削減できることが確認できた。特に、学習に必要な共有を「データ転送」ではなく「モデル更新」に置換することで、通信量削減とデータ保護を同時に達成できる点が重要である。観光地のように上り帯域や回線コストが制約となりやすい環境では、通信量の削減はそのまま導入可能地点の拡大(屋外・屋内等)に繋がる。一方、モデル更新頻度を高めると下り配布や上り更新が増えるため、今後は「更新頻度(例:日次/週次)」と「精度・追従性」のトレードオフを評価し、最適な運用条件を確立する必要がある。

2) バーチャルセンサーの確認

特定地点のエッジデバイスを AI エージェントから切り離し(当該地点の実測データを利用しない状態)としても、他地点データと外部要因、及び学習済み予測モデルにより、当該地点の混雑をバーチャルセンサー(推定値)として出力できることを確認する。これにより、物理デバイスの設置数を削減しつつカバレッジを確保できるかを評価した。

評価条件(切り離し試験)

- ・対象地点:[地点 A/地点 B/…](代表的な混雑特性を持つ地点を選定)
- ・切り離し方法:対象地点の実測混雑入力を停止(学習・推論から除外)
- ・正解データ:評価用に対象地点の実測値は保持し、推定値との誤差を算出

評価指標

- ・推定精度:MAE/RMSE(人数)、混雑カテゴリ一致率(%)
- ・ピーク追従:ピーク時の誤差、最大値再現率、遅れ(分)
- ・運用性:切り離し中もエージェントが回答可能である割合(%)
- ・デバイス削減率:削減可能と判定できた地点数/全地点数(%)

評価結果は表 6.6-3 の通りである。バーチャルセンサー評価により、特定地点の実測データが取得できない(または設置しない)場合でも、他地点の観測データや外部要因を用いて混雑を推定し、エージェントが情報提供を継続できる可能性が示された。これは、設置・保守コストの削減だけでなく、通信設備や設置制約のある地点への適用範囲を拡張する上で有効である。一方で、バーチャルセンサーの精度は、(1) 対象地点と他地点の相関、(2) 流入を説明する上流指標(駐車場・バス降車等)の有無、(3) イベント等の突発要因の反映、に依存する。特にピーク時の追従性が課題となる地点については、上流地点の観測追加や外部データ(イベント・交通)の取り込みにより改善余地がある。

表 6.6-3 対象地点毎の評価結果

| 対象地点 | 混雑カテゴリ一致率 | 判定(代替可否) |
|-----------|-----------|----------|
| 五稜郭タワー入口 | 80% | 不可 |
| 五稜郭タワー出口 | 85% | 代替可 |
| 五稜郭タワー展望台 | 95% | 代替可 |

今後の社会実装においては、「重要地点は実センサー」「補助地点はバーチャル」でカバレッジを確保するハイブリッド構成が合理的であり、地点タイプ別に代替可否の目安(誤差閾値、相関閾値等)を整備することが、他地域展開の指針である。

6.6.3 KPI/KGI との比較結果

| 定性評価 /定量評価 | 番号 | 目標値 |
|---------------|----|---|
| 定量評価 | 1 | AI エージェントに組み込むことで、別途定義している通信量の低減が達成できること |
| | 2 | 特定エッジを AI エージェントから切り離し(データを利用しない状態)にしてもバーチャルセンサーとして動作し、デバイス削減率の KPI を達成できることを評価する |

1) 通信量の低減が達成できること

6.6.2「通信負荷・通信量の比較」で示した方式別比較(方式 A:中央集権型、方式 B:提案方式=エッジ推論+連合学習)に基づき、1日あたりの上り通信量を算出した。その結果、提案方式(方式 B)では、中央集権型(方式 A)と比較して上り通信量が約 99%削減され、通信量低減の目標を達成した。更に、学習目的の生データ転送については、拠点外への転送を行わずモデル更新情報のみを送受信する運用であるため、学習目的の生データ転送量は 0%であった。以上のことから目標値を達成した。

連合学習の特徴である「データを集約せず、更新情報のみ共有する」設計が、通信量低減に直結することを示している。特に観光地等、回線費・上り帯域が制約となる環境では、通信量削減が導入可能地点の拡大につながり、社会実装の実現性を高める。一方、通信量はモデル更新頻度や再送の影響を受けるため、社会実装に向けては、更新頻度(例:日次/週次)と追従性(ドリフト耐性)の最適化、更新差分の圧縮、夜間バッチ更新等の運用設計を含め、安定条件を確立する必要がある。

2) デバイス削減率の KPI を達成できること

対象地点を切り離れた状態で、他地点データ・外部要因・学習済みモデルにより当該地点の混雑を推定し、実測値との誤差を評価した。その結果、バーチャルセンサーとして代替可能(=運用上許容できる精度で推定可能)と判定された地点数は 2 地点/評価対象 3 地点であり、デバイス削減率は 66% となった。バーチャルセンサーによる代替が可能であることは、設置・保守コストの削減に加え、設置制約(屋内・私有地・景観配慮等)や通信制約のある地点のカバレッジ拡大に寄与する。特に、混雑変動が比較的安定した地点では代替可能性が高く、「重要地点は実センサー」「補助地点はバーチャル」というハイブリッド構成により、合理的に設備を削減できることが示唆される。一方、ピークが突発的に発生する地点や、他地点との相関が弱い地点では推定誤差が大きくなりやすい。このため、社会実装に向けては、(1) 代替可否の判定基準(例:混雑カテゴリー一致率、ピーク誤差、遅れ時間)を明確化し、(2) 上流指標(駐車場・交通・イベント等)の外部データを拡充し、(3) 代替できない地点を優先的に実センサー配置する、といった設計指針を整備する必要がある。

6.6.4 考察

エッジ推論、分散連合型 RAG (F-RAG)、連合学習、AI エージェントを統合することで、観光混雑という実環境課題に対し、通信制約・合意形成・プライバシー配慮を同時に満たす実装形態の有効性を確認した。具体的には、(i) 通信量の大幅削減(約 99%級)、(ii) 学習目的の生データ転送ゼロ、(iii) 準リアルタイムな情報提供(300ms 以内を高い割合で達成)、(iv) バーチャルセンサーによるデバイス削減可能性が示された。

本方式は、通信制約のある観光地でも安定運用できることから、複数観光地・施設への横展開が見込まれる。生データを外部に出さずに予測や助言を提供できるため、自治体・施設・事業者間での合意形成が進みやすく、導入障壁が低い。今後は、交通・イベント・気象等の外部データ連携や、サイネージ・アプリとの連動により、混雑回避行動の促進まで含めた価値提供へ発展させることが期待される。一方で、設置環境差による品質揺らぎや突発ピーク対応、運用標準化が課題であり、品質フラグや更新ルールを整備することが重要となる。

中央集権型(画像・データ集約)と比較すると、本方式は推論をエッジで実行し、学習共有をモデル更新に置換することで、通信量を約 99%級まで削減できる。これにより、回線費負担の軽減、回線条件が厳しい地点への導入可能性向上、データ提供に対する心理的・制度的ハードル低減が同時に得られる。また、リアルタイム性は通信速度そのものではなく、現地処理、軽量 API 取得、巡回優先度制御、キャッシュやフォールバックといったシステム設計によって確保できることが示唆された。

本ソリューションは観光分野にとどまらず、商業施設、公共施設、交通、防災等、現場分散型データを扱う多様な領域に適用可能である。「データを動かさずに知見を統合する」構成は、データ連携が進みにくい国内環境と親和性が高く、地域全体のデータ利活用を促進する。ビジネス面では、エッジ運用、通信、クラウド、API、ガバナンス支援等、多層的なサービス展開が期待できる。

また、エッジ推論、連合学習、F-RAG、AI エージェントを統合し、低通信・データ保護・準リアルタイム・対話 UI を同時に成立させた点に先進性がある。実証からは、①画像は送らず数値化する、②連合学習は軽量な時系列モデルから始める、③リアルタイム性は巡回・キャッシュ設計で担保する、④品質揺らぎを前提にルールを標準化する、⑤実センサーとバーチャルセンサーのハイブリッド配置が合理的、という再利用可能な設計知見が得られた。

7. 本実証の総括

7.1 本実証の成果・課題

本実証では、エッジ AI、連合学習、AI エージェント、クエリ先選択型 F-RAG を組み合わせた分散型 AI インフラ基盤について、地域の実環境下での有効性を検証した。その結果、通信負荷の低減、プライバシー配慮、リアルタイム性と予測精度の両立が可能であることが確認され、観光分野における意思決定支援への有用性が示された。一方で、分散型構成に適した運用ルールや費用対効果の定量化、社会受容性確保といった課題も明らかとなった。今後は、段階導入や分野横断展開を通じて、公共領域における持続可能な AI×通信インフラの社会実装の取り組みを継続していく。

■本実証を通じた成果

人流データの取得からエッジ AI による解析、クラウドでの蓄積・提供、更に連合学習、AI エージェント、F-RAG を用いた利活用までを一連の仕組みとして統合し、観光地・市街地という実環境で検証を行った。その結果、通信制約やプライバシー制約が大きい地域環境においても、分散型アーキテクチャが実用的に成立することを確認した。

技術面では、映像解析をエッジ側で完結させ、クラウドへは人数・混雑度等の数値データ及びモデル更新情報のみを送信する構成としたことで、中央集権型(画像・動画送信)と比較して、1日あたりの通信量を最大 99.9%削減できることを定量的に確認した。具体的には、標準的な JPEG 画像転送(約 3.3GB/日)と比較して、時系列混雑予測モデルの連合学習構成では約 0.7MB/日まで通信量を低減でき、モバイル回線を前提とした運用でも成立することが示された。また、イベント駆動通信及び AI エージェントによる巡回制御を組み合わせることで、不要な定期送信を抑制し、通信コスト削減率 20%以上という KPI を安定的に達成できる見通しを得た。通信経路についても、LAN、VPN、暗号化通信を用いた閉域構成とすることで、デバイスからクラウドまで一貫したセキュリティを確保した。リアルタイム性の観点では、ウェブインターフェイスを公開し、観光客利用を想定した検証を実施した結果、AI エージェントへの問い合わせから応答までのエンドツーエンド遅延の 98%が 300ms 以内で処理されていることを確認した。これは、観光客が体感的な遅延を感じることなく混雑情報や予測結果を取得できる水準であり、即時情報提供という目的に対して十分な性能を有している。

AI 面では、連合学習を活用することで、映像データを中央に集約せずとも、各観光スポットや交通拠点の特性を反映したモデル更新が可能であることを確認した。これにより、季節性、時間帯、イベント等の変動を捉えながら、予測精度についてもベースライン手法と比較して 20%以上の改善が見込めることを確認し、高精度な未来予測の実現性を示した。更に、特定のエッジデバイスを意図的に切り離した状態で検証を行い、他地点データと学習済みモデルを用いて混雑度を推定する「バーチャルセンサー」が実用的に機能することを確認した。これにより、全地点への実センサー設置を前提としない構成が可能となり、設置・保守コストを抑えつつエリアカバレッジを確保できる可能性が示された。

サービス・ユースケースの観点では、混雑実測値及び予測結果を AI エージェントが対話的に提示することで、観光客や事業者の行動判断(訪問時間の調整、代替候補の検討等)に資する情報提供が可能であることを確認した。F-RAG を用いることで、最新値は API から、根拠情報は文書データから参照する構成と

し、回答の安定性と説明可能性を両立した。

加えて、行政部局へのヒアリングを通じて、観光分野に加え、防災、都市計画、交通といった分野においても、通行量調査の省力化や避難計画策定への活用等、具体的な導入ニーズが存在することを確認できた点は、社会実装に向けた重要な成果である。

■今後の課題

一方で、実運用・横展開を前提とすると、以下の課題が明確となった。

第一に、運用設計・ガバナンスの具体化である。分散型構成では、エッジ機器の監視・保守、障害時の一次切り分け、モデル更新のタイミングや承認プロセス、データの保存期間・閲覧権限等、運用ルールの整備が不可欠となる。第二に、費用対効果の定量化と予算化の道筋である。通信削減・業務省力化・混雑緩和・満足度向上といった効果を、運用主体が意思決定できる形で指標化する必要がある。導入費(機器・工事・SI・モデル)と運用費(通信・クラウド・保守・監視)を整理した上で、便益(人件費削減、調査代替、誘導効果、リスク低減等)を定量評価し、費用負担の考え方(行政、事業者、共同利用)を明確にすることが重要となる。第三に、社会受容性の継続的確保である。映像をクラウドに送らない、解析結果のみを扱う、暗号化・閉域化を徹底する等の設計は、受容性向上に寄与する一方、住民・関係者に対して「何を取得し、何を取得しないか」「誰が何に使うのか」を平易に説明し続けることが不可欠である。設置場所の合意形成、景観・安全面の配慮、問い合わせ対応の体制整備等、技術以外の運用要件も社会実装の成否を左右する。

■今後の展開方針

これらの成果と課題を踏まえ、今後は本実証で構築した AI インフラ基盤を、段階的かつ分野横断的に展開すべく、函館市と引き続き協議を継続する。具体的には、まず観光・交通分野における実運用を想定した継続利用を進めると共に、防災・都市計画分野への適用を通じて、平時と非常時の双方で価値を発揮する基盤としての検証を深める。また、複数分野での共同利用やプラットフォーム化を進めることで、利用主体の拡大によるスケールメリットを創出し、コスト効率と持続性の向上を図る。並行して、他地域への横展開や他のスマートシティ施策との連携を視野に入れ、地域に根付いた持続可能インフラとして社会実装を進めていくことを目指す。

8. 参考資料

8.1 付録1 F-RAG WebAPI 仕様 詳細

1. 概要

| 項目 | 内容 |
|---------|---|
| エンドポイント | /v1/chat |
| 説明 | ユーザーからのメッセージ一覧を受け取り、AI アシスタントの返信を返すチャット API。session_id により会話の文脈を維持可能。 |
| 利用プロトコル | HTTPS |
| 認証方式 | Bearer Token (※実証では固定 Token) |
| レスポンス形式 | application/json |
| タイムアウト | 30 秒(想定) |
| レート制限 | 必要に応じて設定(例:60 リクエスト/分) |
| 備考 | <ul style="list-style-type: none">・session_id 未指定時はサーバ側で新規作成し、レスポンスで返却・無効な session_id 指定時は 404・stream=true の場合はストリーミング(SSE)で返却(202) |

2. リクエスト

| パラメータ名(論理名) | 物理名 | データ型 | 必須 | 説明 |
|---------------|---------------|--------------------|----|---|
| セッション ID | session_id | string | 任意 | 会話を識別する ID。未指定なら自動新規作成。 |
| メッセージ一覧 | messages | array<object> > | 必須 | <ul style="list-style-type: none">・ユーザーの今回の発話を含むメッセージ配列・API 側でシステムメッセージが 0 番目に挟まれる・messages の型の仕様はレスポンス側を参照 |
| ストリーミングフラグ | stream | boolean | 任意 | true の場合は SSE で返却(202)。デフォルト false。 |
| テナント ID (ヘッダ) | X-Tenant-Id | string | 必須 | 呼び出し元のテナント識別子。 |
| 認証トークン (ヘッダ) | Authorization | string | 必須 | Bearer Token |

3. レスポンス

| パラメータ名 | 物理名 | データ型 | 例 | 説明 |
|------------|------------|--------|---|--|
| セッション ID | session_id | string | "session_id": "ab12345c-67d8-9012-efg3-456789hij0" | 利用中の session_id |
| AI 応答メッセージ | message | object | "messages": [{"role": "user", "content": "こんにちは"}, {"role": "assistant", "content": "こんにちはは!..."}, {"role": "user", "content": "おすすめは?"}, {"role": "assistant", "content": "函館には..."}] | ・会話の Message オブジェクト ・最新の応答をとる場合は、messages[-1]["content"]を指定 |

8.2 付録2 観光客向けアプリケーション動作イメージ

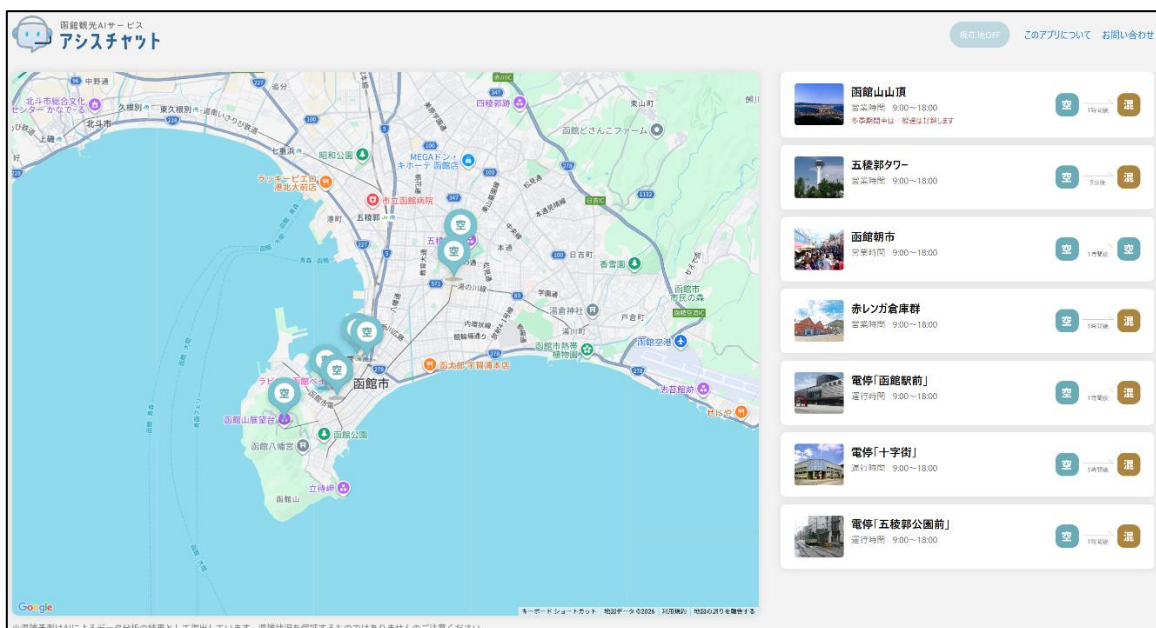


図 8.2-1 初期画面



図 8.2-2 「五稜郭タワー」選択後の画面



図 8.2-3 「五稜郭タワー・#30分間でできることは?」選択後の画面

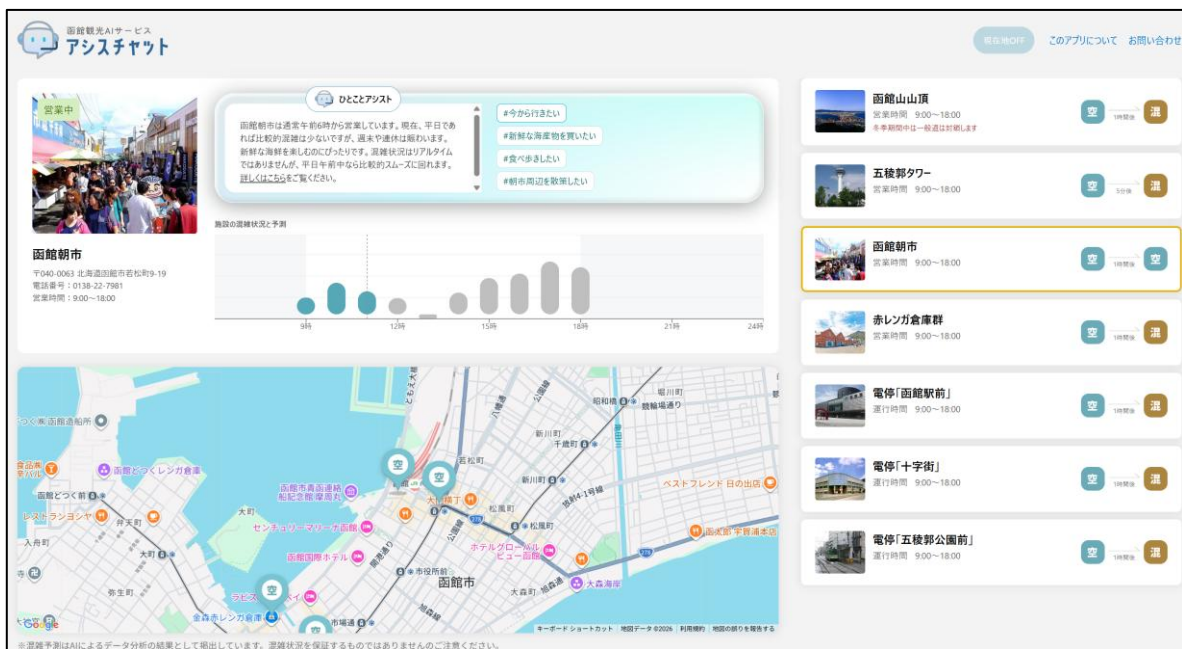


図 8.2-4 「函館朝市」選択後の画面

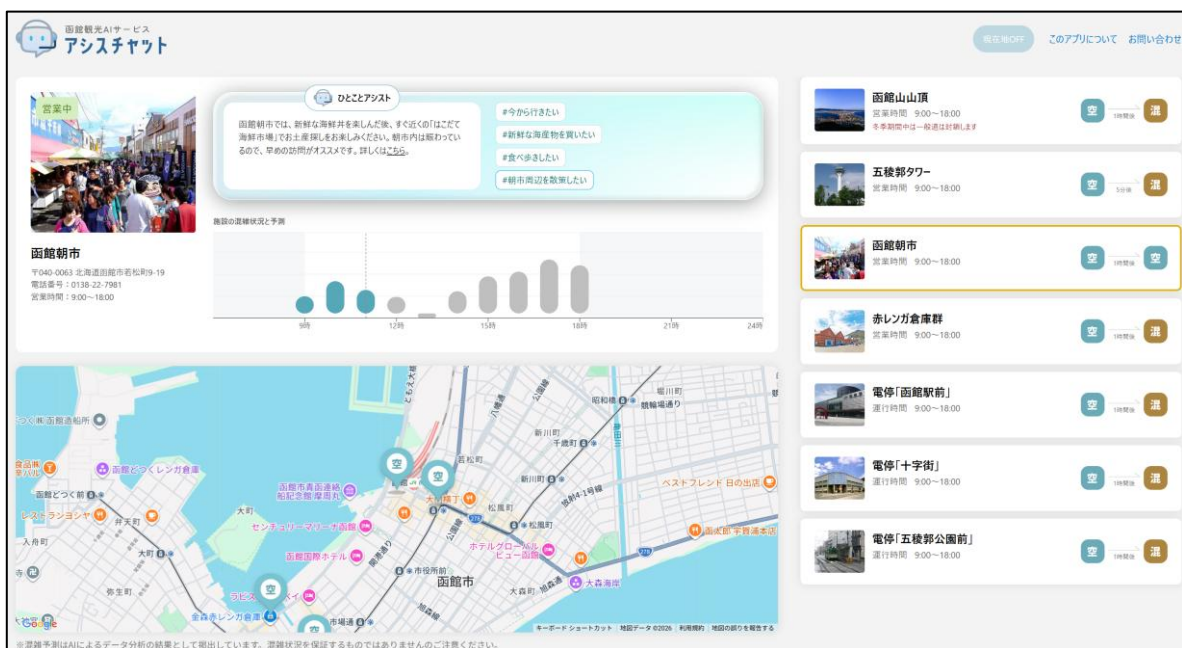


図 8.2-5 「函館朝市・#朝市周辺を散策したい」選択後の画面

8.3 付録3 機械学習モデルの詳細構成(TCN + Node-wise Self-Attention)

本実証で用いた混雑(来訪者数)予測モデルについて、入力層から出力層までの構造、各層の設定パラメータ、学習設定、連合学習(FedAvg)における運用設定を整理する。本文 5.4.5「連合学習の実装」で述べた通り、時間的特徴抽出(TCN)と観光地(ノード)間の相互作用学習(Node-wise Self-Attention)を組み合わせ、軽量かつ実運用可能な予測モデルとして設計した。

モデルの目的・設計方針

各観測地点(ノード)における来訪者数(人数カウント)を、準リアルタイムに取得できる実測値および外部要因(天候・曜日等)から、近未来(P1/P2/P3)を推定する。設計方針として、①時系列の長期依存を扱えること(TCN)、②複数地点の影響関係を取り込めること(Self-Attention)、③連合学習で更新可能な軽量性(パラメータ共有量を抑制)を同時に満たす。

入出力仕様(テンソル形状)

- ・ノード数: $N=15$ (本文「観測地点(ノード)」に列挙した地点)
- ・時間窓: 過去 $L=240$ 分(1分粒度、 $T=240$ ステップ)・入力特徴量数: $F=19$ (詳細は付録4参照)
- ・予測ホライズン: $H=3$ ($P1=60$ 分後、 $P2=120$ 分後、 $P3=240$ 分後)
- ・入力テンソル: $X \in R^{B \times N \times T \times F}$ (B : バッチサイズ)
- ・出力テンソル: $\hat{Y} \in R^{B \times N \times H}$ (各ノードの将来人数予測)

ネットワーク構成(入力層～出力層)

本モデルは、(1) TCN Encoder 1 による時間的特徴抽出、(2) Node-wise Multi-Head Self-Attention によるノード間相互作用の学習、(3) TCN Encoder 2 による時系列上の平滑化・補正、(4) 全結合層による多ホライズン回帰、の4段で構成される。

| 層(ブロック) | 処理内容 | 主な設定パラメータ | 入力形状 | 出力形状 |
|--------------------------|--|---|---------------------------------|---------------------------------|
| Input | 各ノードの時系列特徴量を入力 | $T=240, F=19,$ $N=15$ | $B \times N \times T \times F$ | $B \times N \times T \times F$ |
| TCN Encoder 1 | 因果(Causal) Conv1D+ Residual Block で時間依存を抽出 | Residual Block $\times 4,$ channels=64, kernel=3, dilation={1,2,4,8}, dropout=0.1 | $B \times N \times T \times F$ | $B \times N \times T \times 64$ |
| Node-wise Self-Attention | 各時刻 t でノード間の影響度を学習(Multi-Head) | $d_{model}=64,$ heads=4, dropout=0.1, LayerNorm | $B \times N \times T \times 64$ | $B \times N \times T \times 64$ |

| | | | | |
|------------------|---|--|---------------------------------|---------------------------------|
| TCN Encoder 2 | Attention 後特徴の時間方向の補正・平滑化 | Residual Block×2, channels=32, kernel=3, dilation={1,2}, dropout=0.1 | $B \times N \times T \times 64$ | $B \times N \times T \times 32$ |
| Temporal Pooling | 最終時刻(t=T)特徴を抽出(または平均) | pool=last_step | $B \times N \times T \times 32$ | $B \times N \times 32$ |
| Concatenation | Encoder1 最終特徴(64)と Encoder2 最終特徴(32)を連結(Skip) | concat dim=96 | $B \times N \times (64+32)$ | $B \times N \times 96$ |
| FC(回帰) | 多ホライズン回帰 | Linear 96→64 (ReLU, dropout=0.1)→ Linear 64→3 | $B \times N \times 96$ | $B \times N \times 3$ |

※表中の形状は概念表記であり、実装上は TCN 計算のため $B \times N$ をまとめて $(B \cdot N) \times F \times T$ の形に変換して畳み込みを適用し、その後 $B \times N \times T \times d$ に戻す。尚、Concatenation では、Self-Attention 出力($B \times N \times T \times 64$)から最終時刻の特徴($B \times N \times 64$)を抽出し、TCN Encoder 2 の pool 出力($B \times N \times 32$)と連結して $B \times N \times 96$ の特徴量を生成する。

学習設定(ハイパーパラメータ)

- ・損失関数:MSE(評価指標として RMSE を算出)
- ・最適化手法:Adam(学習率 $1.0e-3$, weight decay $1.0e-5$)
- ・バッチサイズ:128(GPU メモリに応じて 64~256 の範囲で調整)
- ・エポック数:最大 20(Early Stopping:patience=5)
- ・正規化:人数カウントはノード毎に z-score 正規化、連続量(気温・降水量等)も学習用にスケールリング
- ・勾配安定化:Gradient Clipping(L2 ノルム 1.0)を適用

連合学習(Federated Averaging)設定

- ・集約方式:FedAvg(クライアント毎のデータ件数に基づく加重平均)
 - ・ローカル学習:各クライアントで E=1~2 エポックのミニバッチ学習を実施し、モデル重み(または差分)を送信
 - ・ラウンド数:R=10~20 を基本設定(収束確認のため最大 50 ラウンドまで試行)
 - ・通信対象:生データ(映像・個票)は送信せず、モデル更新情報(重みパラメータ)のみを送受信
 - ・更新サイズ目安:本モデルのパラメータは約 0.15~0.20M であり、float32 換算で約 0.6~0.8MB
- 推論時の処理(エッジ/クラウド連携)

推論は、(1) エッジでの人数・混雑度の算出(画像→数値化)、(2) 数値時系列と外部特徴量を用いた予測モデル推論、(3) 予測結果を WebAPI 経由で AI エージェントが参照(F-RAG/MCP)、の順に実行される。予測モデル自体は軽量であるため、エッジ側での推論、または拠点サーバ(クラウド VM)での推論のいずれでも運用可能である。

- ・エッジ推論:最新の時系列(過去 L 分)をローカル DB から抽出し、同一時刻の外部特徴量と結合してモデル入力を生成する。
- ・クラウド推論:集約済みグローバルモデルを予測 API として公開し、AI エージェントが必要な地点・時間帯のみを取得する。

8.4 付録 4 入力特徴量の詳細(入力次元)

混雑(来訪者数)予測モデルに入力する特徴量を、採用条件・前処理を含めて厳密に整理し、入力次元(テンソル形状)が再現できるよう記載する。

前処理と時間窓(入力系列の定義)

- ・基礎データ:エッジ推論により得られる人数カウントを 1 分粒度の時系列として整形する(本文「(2)データセット作成に係る処理」参照)。
- ・入力窓:推論時刻 t に対し、過去 $L=240$ 分($t-239 \sim t$)を入力系列とする($T=240$)。
- ・欠損の扱い:連続欠損が 10 分以上の区間は学習・評価から除外する。連続欠損が 10 分未満の短時間欠損は前値保持または線形補間で補完する。

入力特徴量一覧($F=19$)

| No. | 特徴量名 | 内容 | 符号化(例) | 次元 |
|-----|-----------------------|-----------------|---------------------------|----|
| 1 | count_i(t) | 人数カウント(エッジ推論結果) | 連続値(z-score) | 1 |
| 2 | time_sin(t) | 時刻(1日の周期) | $\sin(2\pi \cdot t/1440)$ | 1 |
| 3 | time_cos(t) | 時刻(1日の周期) | $\cos(2\pi \cdot t/1440)$ | 1 |
| 4 | dow_mon(t)~dow_sun(t) | 曜日 | one-hot(7区分) | 7 |
| 5 | is_holiday(t) | 祝日フラグ | 0/1 | 1 |
| 6 | temp(t) | 気温 | 連続値(標準化) | 1 |
| 7 | precip(t) | 降水量/降雪量 | 連続値(標準化) | 1 |
| 8 | weather_cat(t) | 天気区分 | one-hot(5区分) | 5 |
| 9 | event_flag(t) | イベント有無 | 0/1 | 1 |

- ・上記の外部特徴量(時刻・曜日・天候・イベント)は、同一時刻 t において全ノード共通であるため、モデル入力では各ノードの count_i(t)に結合し、ノード毎に同一の外部特徴量を付与する。
- ・従って、各時刻 t におけるノード i の入力ベクトルは $x_i(t)=[\text{count}_i(t), \text{exog}(t)] \in \mathbb{R}^{\{19\}}$ となる。

入力テンソル形状と教師データ(ラベル)定義

・入力テンソル: $X \in \mathbb{R}^{B \times N \times T \times F} = \mathbb{R}^{B \times 15 \times 240 \times 19}$

・教師データ(予測対象): $Y \in \mathbb{R}^{B \times N \times H} = \mathbb{R}^{B \times 15 \times 3}$

・ラベル定義: 推論時刻 t に対し、 $P1=60$ 分後、 $P2=120$ 分後、 $P3=240$ 分後の人数カウントを目的変数とする。

$$Y_i = [\text{count}_i(t+60), \text{count}_i(t+120), \text{count}_i(t+240)]$$

データ採用条件(品質フラグ・除外条件)

・連続欠損 10 分以上: 学習・評価対象から除外(センサー停止・通信断等で信頼性を担保できないため)

・短時間欠損(10 分未満): 前値保持または線形補間で補完

・外れ値: 物理的に不可能な急増減(例: 急激なゼロ化等)は外れ値検知により除外または補正

・低品質条件: 逆光・遮蔽・夜間等で推論品質が著しく低下する区間は品質フラグに基づき除外(または重みを下げて学習)

地域社会 DX 推進パッケージ事業(AI 検証タイプ)

AI を用いた通信負荷の低減・通信量の確保等の検証に関する調査研究
(実証機関:ソフトバンク株式会社)

実証報告書

観光地における連合学習を用いた多地点混雑予測 AI の学習データ多様化の実証

2026 年 1 月

AI01 函館市
ソフトバンク株式会社
国立大学法人東京大学、宇宙サービスイノベーションラボ事業協同組合
