**논문번호:** J1\_202500002

**원논문 제목:** “설명가능한 인공지능 활용 서울시 지하철 혼잡도 예측과 정책적 의사결정 지원을 위한 비즈니스 애널리틱스(Business Analytics to Predict Seoul Subway Congestion and Policy Decision-making using Explainable AI)”

**안녕하십니까 한국빅데이터학회 학회지 에디터님,**

원고의 재제출과 심사자의 의견에 대해 답변할 수 있는 기회를 주셔서 감사합니다.

요청해주신 바와 같이 아래에 심사자별 코멘트에 대한 답변을 포함하였습니다.

심사자분들께서 지적해 주신 내용 외에도 일부 업데이트 사항을 반영 및 표시하여 논문을 개선하였습니다.

감사합니다,

박재흥, 김경원

## 심사자 #2

### 심사자 #2, 심사내용 #1 & #2

|  |
| --- |
| 선행연구가 절대적으로 부족합니다. 혼잡도 관련된 연구들을 보완하시고, 데이터 방법론 관련된 부분, 정책적 시사점 관련된 논문들을 최소 20편이상을 추가해 주십시오. 기존 연구와의 차별성을 충분히 논의할 필요가 있음.  특히 해외 관련 사례들 연구도 추가하여 우리나라 정책들과 비교분석도 결론에서 추진할 필요가 있습니다.  1) 지하철 혼잡도 예측 관련 선행연구  2) 머신러닝 및 딥러닝 기법을 활용한 교통 혼잡도 예측 연구  3) 인공지능(XAI)의 교통 분야 적용 연구  연구의 차별성 및 기여도를 보다 명확히 제시해 주십시오.  1. 6개의 공공데이터베이스를 융합한 접근법이 어떻게 기존 연구와 차별성  2. 95% 이상의 예측 정확성이 기존 연구들(예: Jang 외(2023)의 80% 정확도)과 비교, 이의 의의  3. SHAP를 활용한 설명 가능한 인공지능 접근법이 기존 블랙박스 모델과의 차별성 |

**저자답변:** 본 연구의 기여점을 강조할 수 있도록 중요한 코멘트를 주셔서 감사합니다. 기존의 지하철 혼잡도 예측이 다소 전통적이며 제한적인 데이터와 알고리즘을 사용하고 있었고 설명가능한 인공지능을 활용한 해석에 관련된 연구는 전무한 상황이었습니다. 그리고 기존 연구들에서 얼마나 제한적인 데이터와 알고리즘을 사용했는지 그리고 성능은 어떠했는지 등을 <표 1>에 반영함으로써 기존 연구 대비 본 연구의 기여를 독자들이 쉽게 이해할 수 있게 되었습니다. 아울러 SHAP 알고리즘의 설명을 추가하여 필요성과 차별성을 강조할 수 있었습니다. 감사합니다.

**추가사항:**

|  |
| --- |
| 또한 지하철 뿐만 아니라 버스나 교통상황들을 예측하기 위해 도메인 특성에 맞는 딥러닝 알고리즘들을 개발하여 높은 성능을 달성한 연구들도 많아지고 있다 [20-21]. 그 외에 관련 선행연구는 <표 1>의 연구목적 별로 상세하게 확인할 수 있다. 전반적으로 머신러닝과 딥러닝 알고리즘을 활용한 혼잡도 예측 연구가 성능이 증가되고 있음을 알 수 있다.  인공지능의 대표적 알고리즘인 머신러닝과 딥러닝은 높은 성능의 예측값을 제공하지만, 예측의 근거를 사람이 직관적으로 이해하기 어렵다는 한계가 있다. 특히 머신러닝과 딥러닝으로 진화하면서 “블랙박스(Black Box)” 모델로 불릴 정도로 내부 의사결정 과정이 불투명해, 사용자가 모델의 결정 이유를 해석하거나 검증하기 어렵다. 이러한 한계를 극복하고자 최근 설명 가능한 인공지능(Explainable AI, XAI) 에 대한 관심이 높아지고 있다. XAI는 머신러닝, 딥러닝 등 복잡한 인공지능 알고리즘의 결과를 사람이 이해할 수 있는 형태로 설명하는 기술이다 [27]. 단순히 “무엇을 예측했는가”를 넘어 “왜 그렇게 예측했는가”를 설명할 수 있게 하며, 모델 개발자 뿐만 아니라 최종 사용자, 정책 입안자, 서비스 관리자에게 중요한 정보를 제공한다. 이를 통해 사용자는 모델이 어떤 특성과 패턴을 기반으로 결정을 내렸는지 알 수 있으며 모델의 신뢰성(Trustworthiness)과 투명성(Transparency)를 높이는데 기여한다 [28]. 지하철 혼잡도와 관련되서는 아직까지 XAI를 활용한 연구들이 전무하지만, 다른 교통 분야에서는 인공지능 기반의 높은 성능 뿐만 아니라 해석까지 포함하고 있는 연구들이 늘어나고 있다. 특히 버스의 승객수를 예측하거나 [22], 인구밀도 [23]에 활용될 뿐만 아니라 교통상황과 사고를 상세하게 이해하기 위한 노력들이 증가하고 있다 [24-26].  하지만 XAI를 활용하기 위해선 분석 모델링 이후 별도의 추가 계산이 필요하고 방법론에 따라서는 전체(Global) 설명보다는 부분적인 설명에 그치며 설명 결과가 비전문가에게는 다소 어려울 수 있다. 또한 모든 데이터와 모델에 완벽히 적용되지는 않으며, 복잡한 상호작용에 대한 설명과 실무 적용에는 여전히 기술적 도전이 요구되는 상황이다. 그럼에도 불구하고 이러한 연구들이 확산되고 여러 연구자들에게 이해도를 높일 수 있는 연구들이 많이 필요하다. 대표적인 XAI 기법에는 반사실적 설명(Counterfactual Explanations), 변수 중요도(Feature Importance), LIME (Local Interpretable Model-agnostic Explanations), SHAP (Shapley Additive Explanation) 등이 있다. 본 연구에서는 이러한 방법들 중 가장 기술적으로 완성도가 높은 SHAP (Shapley Additive explanation) 방법을 이용하여 고성능의 지하철 혼잡도 예측의 근거를 확인하고 변수들의 기여도를 분석한다. 게임 이론 기반의 세플리 값(Shapley Value)을 활용해 각 변수의 기여도를 계산하므로 직관적이고 일관된 설명을 제공한다. 지하철 혼잡도 예측을 위해 기존에는 정성적인 접근 또는 제한된 인공지능 알고리즘을 활용하였지만, 본 연구는 대표적인 머신러닝과 딥러닝을 모두 활용하는 예측 방법론을 제공하는 동시에 설명가능한 인공지능을 통해 지하철 혼잡도 대응을 위한 접근을 추가함으로써 기존 문헌들을 확장하였다. |

### 심사자 #2, 심사내용 #3

|  |
| --- |
| 결과 해석과 정책 제언에 있어서 해외 사례와의 비교 보완이 필요함. |

**저자답변:** 결론부분에 정책적인 접근과 효과와 관련된 선행연구를 추가하고 본 연구가 어떠한 트랜드를 반영하고 있는지를 강조 함으로써 연구의 기여도를 더욱 잘 설명할 수 있는 계기가 되었습니다. 감사합니다.

**수정사항:**

|  |  |
| --- | --- |
| [기존]  최근 전 세계적으로 확산하였던 코로나바이러스 역시 사람들이 밀집되는 폐쇄적 대중교통에서 강한 전염성을 보여 지하철 이용이 크게 감소하였기 때문에, 의학적으로나 서비스적으로 혼잡도 이슈는 다양한 방향으로 확산하기 쉽다. 사회적 이슈들을 개선하기 위해서는 다수가 이용하는 대표적 교통수단인 지하철의 혼잡도를 사전에 예측하는 것은 서비스의 효율성과 그리고 시민들의 편의성과 안정성 측면에서 중요한 의의가 있다. | [변경]  최근 전 세계적으로 확산하였던 코로나바이러스 역시 사람들이 밀집되는 폐쇄적 대중교통에서 강한 전염성을 보여 지하철 이용이 크게 감소하였기 때문에, 의학적으로나 서비스적으로 혼잡도 이슈는 다양한 방향으로 확산하기 쉽다. 사회적 이슈들을 개선하기 위해서는 다수가 이용하는 대표적 교통수단인 지하철의 혼잡도를 사전에 예측하는 것은 서비스의 효율성과 그리고 시민들의 편의성과 안정성 측면에서 중요한 의의가 있다. 한국의 대중교통들 중 지하철은 매일 약 800만 명이 이용할 만큼 가장 대표적인 교통수단이기 때문에, 지하철을 이용하는 시민들이 느끼는 삶의 만족도와 관련이 없을 수 없다. 특히 수도권에 인구의 절반이 분포된 한국의 경우 인구밀도의 문제가 일반적인 거리에서의 사고뿐만 아니라 교통수단 내외에서 발생할 수 있는 사고들과도 관련성이 높을 수 있으므로 혼잡도를 낮추기 위해 정책적으로 많은 논의가 이루어지고 있다. 따라서 지하철을 이용하는 시민이라면 역사를 확장하거나 리모델링을 하며 환승 시설의 확장과 개선으로 승객의 흐름을 분산하려는 모습, 출퇴근이나 사회적 이슈에 따른 밀집도를 낮추기 위해 배차 간격을 조정하거나 환승 교통수단들을 확장하여 이용 수요를 분산하려는 노력을 종종 볼 수 있다. 또한, 이용 시간대에 따라 요금을 조정하는 등의 방안들도 정책적으로 제시되곤 있지만, 한국은 아직까진 반영하진 않고 있다. 실제 이러한 노력은 전 세계적으로 일반적으로 나타나고 있는 모습이다. 중국은 베이징의 과거 지하철 노선이 확장되면서 운행 속도가 4% 증가되고 혼잡도가 급격히 감소하였다 [29]. 그리고 혼잡통행료, 대중교통 확장, 자전거 도로 확장 등의 정책도 교통 혼잡과 대기오염 감소에 긍정적임을 확인하였다 [30]. 또한 런던에서는 2003년부터 중심부에 진입하는 차량들에 대해 혼잡통행료를 부과하여 교통량을 10% 감소시키고 배출가스를 줄였으며 교통사고를 40%까지 감소시켰다. 이러한 정책은 뉴욕, 싱가포르, 스톡홀름 등 다른 도시에도 영향을 주고 있다 [31]. 그 외에도 여러 문헌조사와 37명의 전문가 인터뷰 등을 통해 혼잡도 완화를 위한 정책과 전략들을 확인 할 수 있었다 [32]. 하지만 전반적으로 교통공학, 도시계획, 교통정책 등을 중심으로 진행되었고, 최근에는 빅데이터와 AI 기반 최적화를 통해 실시간 운행 스케줄이나 혼잡도 예측하여 알려 줌으로써 승객의 편의를 증진시킴으로써 자연스러운 승객의 분산을 유도하려는 노력들이 지속되고 있다. |

⬛