

ChatGPT for Research and Publication: A Step-by-Step Guides

Reviewer : Seokhun Park



Department of Statistics
Seoul National University

Contents

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

- 이 논문은 Chatgpt를 연구와 논문작성 과정에 활용하기 위한 단계별 가이드를 제공함.
 - ① 논문 제목 생성
 - ② 초록 작성
 - ③ 서론 작성
 - ④ 선행연구 및 연구동향 작성
 - ⑤ 방법론 구상
 - ⑥ 참고문헌 작성
- 현재 사용하고 있는 최신 버전의 chatgpt 4o 를 이용해서, 해당 논문에서 제시한 가이드를 따라 실험을 진행해봄.

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

제목 생성

- ChatGPT가 연구 주제, 목적 등을 입력받아 여러 논문 제목 후보를 제안해 줄 수 있음.

Functional ANOVA model의 component를 Tensor Product Neural Network (TPNN)으로 추정하는 모형을 제안하는 논문의 제목을 지어줘

🔥 다양한 스타일의 제목 제안

◆ 방법론 제안 중심

1. "A Tensor Product Neural Network Model for Functional ANOVA Decomposition"
2. "Estimating Functional ANOVA Components via Tensor Product Neural Networks"
3. "TPNN: A Neural Architecture for Functional ANOVA Component Estimation"

◆ 약간 학술적이고 포괄한 느낌

4. "Nonparametric Estimation of Functional ANOVA Components Using Tensor Product Neural Networks"
5. "A Deep Tensor Product Network for Learning Functional ANOVA Structures"

◆ 간결하고 임팩트 있는 스타일

6. "Neural Estimation of Functional ANOVA via Tensor Product Networks"
7. "Functional ANOVA Modeling with TPNNs"
8. "TPNN: Learning Additive Structure with Functional ANOVA"

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

- 연구의 주요 아이디어나 전체 본문을 ChatGPT에게 입력하면, 그것을 요약해 짧은 초록 형식의 텍스트를 생성할 수 있음.
- 논문의 제목만으로 초록을 생성해줄수 있지만, 더 많은 정보를 제공할수록 Chatgpt는 더 정확한 정보가 담긴 초록을 생성해줌.

- ANOVA-TPNN논문의 introduction 을 바탕으로 논문 초록을 작성해달라고 했을때의 결과.

As artificial intelligence (AI) models have become increasingly complex, interpretability has emerged as a key concern. While post-hoc explanation methods such as LIME and SHAP have been widely used, they often provide only local approximations of black-box models and fail to deliver faithful global interpretations. The functional ANOVA model, which decomposes a high-dimensional function into a sum of low-dimensional components, offers a promising alternative as a transparent-box model for interpretable AI. However, recent neural implementations of this model, such as NAM and NBM, suffer from unidentifiability issues, making component-wise interpretation unreliable.

In this paper, we propose a new neural network architecture, **Tensor Product Neural Networks (TPNN)**, for the functional ANOVA model. TPNN leverages tensor product basis expansion and replaces traditional basis functions with specially designed neural networks to ensure identifiability and robustness. This design allows each component to be accurately and stably learned via standard stochastic gradient descent algorithms. We further introduce **ANOVA-TPNN**, a model that integrates TPNNs to estimate the entire functional ANOVA decomposition.

We theoretically establish the universal approximation property of ANOVA-TPNN for Lipschitz functions. Empirical results on multiple benchmark datasets demonstrate that our method provides more accurate, stable, and interpretable component estimates than existing models such as NAM, NBM, NODE-GAM, and XGBoost—without compromising predictive performance. Furthermore, our model supports practical constraints such as outlier robustness and monotonicity, making it suitable for trustworthy AI applications.

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성**
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

서론 (Introduction) 작성

- 연구 배경, 중요성, 목적 등을 정리해 입력하면, hatGPT를 활용해 서론 초안을 작성하는 데 도움을 받을 수 있음.
- 단, 선행연구에 대한 조사는 잘 하지 못함.

설명가능 모형인 Functional ANOVA model 의 component를 TPNN으로 추정하는 모형을 제안하는 논문의 서론을 영어로 작성해줘

Chatgpt : In recent years, the demand for interpretable machine learning models has grown alongside the rapid advancement of highly expressive, yet often opaque, predictive algorithms. While deep neural networks and other complex models have achieved impressive accuracy across a wide range of tasks, their lack of interpretability poses significant challenges in high-stakes domains such as healthcare, finance, and scientific discovery. In response to these concerns, functional decomposition techniques, such as the *Functional ANOVA* (fANOVA) model, have gained attention for their ability to provide structured, interpretable representations of multivariate functions.

The Functional ANOVA model decomposes a multivariate function into a sum of components representing main effects and higher-order interactions among input variables. This decomposition not only enhances interpretability by isolating the contribution of individual features and their combinations, but also supports model simplification, variable selection, and sensitivity analysis. However, conventional approaches to estimating fANOVA components—such as kernel methods, additive models, or basis function expansions—often suffer from limited flexibility, scalability issues, or challenges in capturing complex nonlinear interactions, especially in high-dimensional settings.

서론 (Introduction) 작성

Chatgpt : To address these limitations, we propose a novel approach that leverages the expressiveness of neural networks while maintaining interpretability through a structured architecture. Specifically, we introduce a model that estimates the components of the Functional ANOVA decomposition using a **Tree-structured Polynomial Neural Network (TPNN)**. The TPNN architecture is designed to represent multivariate polynomial functions in a hierarchical and interpretable manner, enabling the model to capture intricate interactions between variables while preserving the decomposability inherent to the fANOVA framework.

Our proposed method bridges the gap between interpretability and expressiveness, offering a new paradigm for function approximation that is both transparent and powerful. We demonstrate that the TPNN-based fANOVA model can effectively recover underlying functional structures and provide meaningful insights into variable interactions, while maintaining competitive predictive performance.

- TPNN에 대한 정보를 안주었더니, TPNN = Tree-Structured Polynomial Neural Network 라고 마음대로 설정함.
- 이와 같이, 정보를 부족하게 주면 잘못된 내용으로 작성할 가능성이 있음.

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

- ChatGPT가 주제와 관련된 기본 정보나 연구동향을 나열하는 데 쓸 수 있지만, 최신 논문 혹은 특정 저널들에 대해서 접근 권한 등 한계가 있음에 주의해야 함.

basis expansion을 이용하는 방법으로 functional ANOVA model의 component를 추정하는 최신논문들을 찾고 설명해줘

Chatgpt : 1. Tensor Product Neural Networks for Functional ANOVA Model (2025년 2월)

이 논문에서는 functional ANOVA 모델에서 각 구성 요소를 안정적으로 추정하기 위한 새로운 신경망인 ANOVA Tensor Product Neural Network (ANOVA-TPNN)을 제안합니다. 기존의 신경망은 구성 요소의 비식별성 문제로 인해 추정이 불안정했으나, ANOVA-TPNN은 텐서 곱 기반의 확장을 통해 이러한 문제를 해결합니다. 이 접근법은 모든 매끄러운 함수를 잘 근사할 수 있음을 이론적으로 증명하며, 실험을 통해 기존 모델보다 구성 요소의 추정이 더 안정적이고 정확함을 보여줍니다. [arXiv](#)

2. Basis Expansion Approaches for Functional Analysis of Variance with Repeated Measures (2024년 2월)

이 연구는 반복 측정된 함수형 데이터를 분석하기 위한 새로운 FANOVA 접근법을 제시합니다. 각 샘플 곡선을 basis expansion으로 표현하여, 이차 FANOVA 문제를 다변량 ANOVA 문제로 변환합니다. 이 방법은 완전한 곡선 정보를 활용하여 그룹 간 평균 곡선의 차이를 검정하며, 시뮬레이션과 실제 보행 데이터에 적용하여 그 성능을 입증합니다. [arXiv](#)

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상**
- 7 참고문헌 작성

- 연구 목적, 실험 설계, 자료수집 방안 등을 ChatGPT와의 상호작용을 통해 함께 구상해볼 수 있음.
- ANOVA-TPNN 논문 초록에서의 설명을 제공하고, ANOVA-TPNN을 확장시킬 수 있는 아이디어를 달라고 요청해봄.
- 결과는 다음과 같음.

Chatgpt : 📌 1. Hierarchical Functional ANOVA로 확장

- 아이디어: ANOVA 구성 요소를 계층적으로 정리해, higher-order interaction term도 의미 있게 추정하도록 확장.
- 예시:
 - 주효과 $f_i(x_i)$, 2차 상호작용 $f_{ij}(x_i, x_j)$, 3차 $f_{ijk}(x_i, x_j, x_k)$ 등을 계층적 제약 하에 학습
- 기대 효과: 고차 상호작용 항의 불필요한 overfitting을 방지하면서도, 설명 가능한 고차 모델 구성 가능

📌 2. 스파스 제약(Sparsity Constraint) 결합

- 아이디어: 일부 구성 요소(주효과/상호작용 효과)는 중요하지 않을 수 있으므로, ℓ_1 penalty 또는 group Lasso를 적용
- 결합 방식:
 - 각 구성 요소 $f_S(x_S)$ 에 대해 학습 시 penalty를 부여하여 필요 없는 효과는 제거
- 기대 효과: 변수 선택, 모델 간결성 향상, 계산 효율성 증가

Chatgpt : 📌 3. Bayesian ANOVA-TPNN

- 아이디어: 각 구성 요소를 사전분포로 두고, 베이지안 추론을 통해 불확실성까지 정량화
 - 기술 스택: Variational Inference, MCMC, Deep Gaussian Process 기반
 - 기대 효과: 구성 요소에 대한 신뢰 구간 추정, 불확실성 기반 의사결정
-

📌 4. 동적/시계열 함수형 ANOVA로의 확장

- 아이디어: 입력이 시계열 또는 시공간 데이터일 경우, 구성 요소도 시간에 따라 변화하도록 모델링
- 예시:
 - $f_i(x_i, t)$, $f_{ij}(x_i, x_j, t)$ 등 시간 종속형 항
- 기술적 고려:
 - Temporal basis (Fourier, B-spline over time)
 - RNN 구조와 결합한 동적 ANOVA-TPNN

Outline

- 1 Introduction
- 2 제목 생성
- 3 초록 작성
- 4 서론 작성
- 5 선행연구 및 연구동향 작성
- 6 방법론 구상
- 7 참고문헌 작성

- ChatGPT가 만들어 내는 참고문헌은 종종 실제로 존재하지 않거나 형식이 잘못된 경우가 많음.
- 따라서, 이미 정리된 참고문헌을 형식을 바꿀때(APA → Chicago) 사용하는 것을 추천함.

기타 - Chatgpt를 활용해서 증명 아이디어 얻기

- Chatgpt o1 (추론모델) 을 활용해서 증명에 대한 아이디어를 얻을 수도 있음.
- Detail한 부분에서는 대부분 틀리거나 알려주지 않지만, 전체적인 증명 방향에 있어 의미 있는 아이디어를 제시해주기도 함.
- Example) ANOVA-TPNN의 approximation lemma 증명.

- 아래와 같이 Chatgpt 1o (추론모델) 에게 질문을 했음.

질문 : Consider L-Lipschitz continuous true function $g_{0,S}$ which satisfies the Sum-to-zero, i.e.,

$$\forall j \in S, \forall \mathbf{z} \in \mathcal{X}_{S \setminus \{j\}}, \int_{\mathcal{X}_j} g_{0,S}(\mathbf{x}_{S \setminus \{j\}} = \mathbf{z}, x_j) \mu_j(dx_j) = 0.$$

Let $\phi_j(x|\theta) = \left\{ 1 - \sigma\left(\frac{x-b}{\gamma}\right) \right\} + c_j(b, \gamma) \sigma\left(\frac{x-b}{\gamma}\right)$, where

$\theta = (b, \gamma), \sigma(x) = 1/(1 + \exp(-x))$ and $c_j(b, \gamma) = -(1 - \eta_j(b, \gamma))/\eta_j(b, \gamma)$ with

$\eta_j(b, \gamma) = \int_{x \in \mathcal{X}_j} \sigma\left(\frac{x-b}{\gamma}\right) \mu_j(dx)$. Let $\phi_S(\mathbf{x}_S | \theta_{S,k}) = \prod_{j \in S} \phi_j(x_j | \theta_{j,k}^S)$.

What is the upper bound of $\left\| g_{0,S}(\cdot) - \sum_{k=1}^{K_S} \beta_{S,k} \phi_S(\cdot | \theta_{S,k}) \right\|_{\infty}$?

Chatgpt : 핵심 아이디어

1. Lipchitz 조건 \Rightarrow 도메인의 작은 블록 근사

$g_{0,S}$ 가 L -Lipchitz이므로, \mathcal{X}_S 를 충분히 많은 (직육면체 형태의) 작은 격자로 나누고, 각 격자(box)마다 함숫값을 상수로 근사했을 때 생기는 최대 오차가 대략 $\mathcal{O}(\frac{1}{M})$ 정도로 줄일 수 있다는 고전적 아이디어가 있습니다.

- M 은 각 좌표축 방향으로 나눈 구간의 개수라고 볼 수 있고, 결과적으로 \mathcal{X}_S 전 영역을 $M^{|S|}$ 개 작은 박스로 나눈 것에 해당합니다.

2. Logistic 시그모이드를 이용한 'Step function' 근사

$\sigma(\frac{x-b}{\gamma})$ 에서 γ 를 매우 작게 하면 계단함수(Heaviside 계단함수)에 가깝게 만들 수 있습니다. 이를 여러 좌표에 대해 곱하면 직사체(하이퍼 박스) 지시함수 같은 모양을 만들 수 있으므로, 적절히 조합하면 \mathcal{X}_S 의 부분 집합(박스) 지시함수를 근사할 수 있습니다.

3. Sum-to-zero 제약

우리가 사용하는 ϕ_j 는 각 좌표 j 에 대해 μ_j 적분값이 0이 되도록 $c_j(b, \gamma)$ 를 정의해 두었습니다. 이는 $g_{0,S}$ 또한 각 좌표 적분값이 0이므로(문제에서의 sum-to-zero 가정), 비슷한 조건을 만족하는 함수들(즉 $\int \phi_S = 0$)의 유한 합으로 $g_{0,S}$ 를 근사하기 위해 필요한 구조적 조건을 갖추게 합니다.

4. 오차 경계치

- 좌표별로 M 등분하면 전체 박스(격자) 수가 $M^{|S|}$ 개가 되고, 이를 근사하기 위한 ϕ_S 항의 수도 대략 $M^{|S|}$ 개(K_S)가 필요합니다.
- 한편, 각 박스 내부에서의 Lipchitz 근사는 박스 한 변의 길이를 $\text{diam}(\mathcal{X}_S)/M$ 정도로 잡으면, 해당 박스 내에서 상수로 근사했을 때 $g_{0,S}$ 가 달라지는 최대 폭이 $L \times (\text{직경}/M)$ 정도로 억제됩니다.
- 결국 $K_S \approx M^{|S|}$ 로 두었을 때, $M \approx (K_S)^{1/|S|}$ 이고, 오차는 $\frac{1}{M} \approx (K_S)^{-1/|S|}$ 정도로 떨어집니다.

Thank You