
EXPLAINING KADINSKY PATTERS CLASSIFICATION

A PREPRINT

Karol Degórski, Zuzanna Glinka, Adrian Kamiński
Faculty of Mathematics and Information Systems
Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw

May 19, 2022

ABSTRACT

TODO

Keywords Kadinsky Patterns · XAI

1 Introduction

- Motivation
 - chęć zmierzenia się z jednym z challenge Kadinsky Patterns (zdefiniowanie własnej wersji)
 - stworzenie własnych wyjaśnień problemu
 - porównanie działania automatycznej metody i stworzonej przez nas
- Related work
 - Kandinsky Patterns
 - A Unified Approach to Interpreting Model Predictions
 - ...

2 Methods

- tworzenie danych, dostęp do pełnej informacji o każdym z kształtów (dokładne położenie, wielkość, kolor, kształt, przynależność do 'dużej' figury) - które to informacje stały się podstawą do dalszej pracy
- porównanie działania Lime, IntegratedGradients i Shap -> ostatecznie Shap
- wyciągnięcie z Shap informacji o importance

3 Results

3.1 Experiment 1

- eksperyment korzystający z metody automatycznej shap

3.2 Experiment 2

- eksperyment korzystający z naszej metody

4 Discussion

- porównanie wyników i wyciągnąć wnioski

5 Conclusion

TODO

References