

# Analiza Porównawcza Grafowych Sieci Neuronowych na Datasecie Cora

Przemysław Dąbrowski

## Wprowadzenie

Przeprowadzona analiza porównuje wydajność trzech architektur grafowych sieci neuronowych (GCN, GAT, GraphSAGE) na datasecie Cora w różnych konfiguracjach. Badanie obejmuje oryginalny dataset oraz warianty z losowymi cechami węzłów o wymiarach 1433 i 64.

## 1. Tabele Wydajności - Kluczowe Wyniki

### 1.1 Wydajność na Oryginalnym Datasecie Cora

Model	Test Accuracy	Val Accuracy	Średnia Precision	Średnia Recall	Średnia F1-Score	Confidence (Mean)
GAT	80.20%	78.00%	0.7818	0.8152	0.7952	0.8667
GraphSAGE	79.80%	77.00%	0.7789	0.8219	0.7935	0.7964
GCN	79.50%	77.00%	0.7714	0.8182	0.7879	0.8386

### 1.2 Wydajność na losowych cechach (dim=1433, seed=42)

Model	Test Accuracy	Val Accuracy	Średnia Precision	Średnia Recall	Średnia F1-Score	Confidence (Mean)
GAT	52.80%	51.20%	0.5378	0.5704	0.5212	0.4671
GCN	51.60%	52.00%	0.5073	0.5621	0.5144	0.7000
GraphSAGE	42.10%	40.80%	0.4128	0.4482	0.4126	0.4354

### 1.3 Wydajność na losowych cechach (dim=64, seed=123)

Model	Test Accuracy	Val Accuracy	Średnia Precision	Średnia Recall	Średnia F1-Score	Confidence (Mean)
GAT	43.60%	36.80%	0.4263	0.4512	0.4154	0.2274
GCN	35.90%	33.20%	0.3480	0.3889	0.3504	0.7288
GraphSAGE	27.10%	24.80%	0.2589	0.2810	0.2582	0.6439

## 2. Kluczowe Odkrycia

### 2.1 Hierarchia Wydajności na Oryginalnym Datasecie

1. **GAT wykazuje najwyższą dokładność (80.20%)** z marginalną przewagą nad GraphSAGE (79.80%) i GCN (79.50%)
2. **Wszystkie modele osiągają podobny poziom wydajności** - różnice nie przekraczają 0.7 punktów procentowych
3. **GAT ma najwyższą pewność predykcji (0.8667)**, co sugeruje stabilność modelu

### 2.2 Dramatyczna Degradacja przy losowości cech

1. **Utrata 25-40 punktów procentowych** dokładności przy randomizacji dim=1433
2. **Jeszcze większa degradacja** przy dim=64 (spadek o 35-50 p.p.)
3. **GAT wykazuje największą odporność** na destrukcję cech
4. **GraphSAGE najbardziej wrażliwy** na utratę cech

### 2.3 Wzorce Konwergencji i Trenowania

#### Oryginalny Dataset:

- **Wszystkie modele osiągają 100% dokładności na zbiorze treningowym** - oznacza to silne overfitting
- **GAT wykazuje najwolniejszą konwergencję** (Loss: 0.3690 w epoce 200)
- **GraphSAGE najszybciej konverguje** (Loss: 0.0025 w epoce 200)
- **GCN osiąga umiarkowaną szybkość konwergencji** (Loss: 0.0396 w epoce 200)

#### Losowe cechy:

- **GAT utrzymuje względnie wysoką dokładność treningową** (97.14% dla dim=1433)
- **GCN i GraphSAGE nadal osiągają 100% na treningowym** mimo słabych wyników na testowym
- **Zwiększona rozbieżność między train/test** wskazuje na problemy z generalizacją

### 3. Analiza Per-Klasa

#### 3.1 Najlepiej Klasyfikowane Kategorie (Oryginalny Dataset)

1. **Neural\_Networks**: 90.28% (GAT), 92.36% (GCN), 86.11% (GraphSAGE)
2. **Genetic\_Algorithms**: 85.71% (GAT), 86.81% (GCN), 92.31% (GraphSAGE)
3. **Theory**: 85.94% (GAT), 89.06% (GCN), 87.50% (GraphSAGE)

#### 3.2 Najtrudniejsze Kategorie do Klasyfikacji

1. **Case\_Based**: 74.62% (wszystkie modele)
2. **Probabilistic\_Methods**: 75.86% (GAT), 71.47% (GCN), 71.79% (GraphSAGE)

#### 3.3 Wpływ Randomizacji na Różne Kategorie

- **Neural\_Networks** utrzymują względnie wysoką **dokładność** nawet po randomizacji
- **Probabilistic\_Methods** wykazują **największą degradację** wydajności
- **Case\_Based** i **Rule\_Learning** są szczególnie wrażliwe na utratę cech węzłów

### 4. Rozkład Pewności Predykcji

#### 4.1 Oryginalny Dataset - Wysoka Pewność

- **GAT**: 74.2% predykcji z wysoką pewnością ( $>0.8$ )
- **GCN**: 70.0% predykcji z wysoką pewnością
- **GraphSAGE**: 62.5% predykcji z wysoką pewnością

#### 4.2 Randomizowane Cechy - Niska Pewność

- **Dim=1433**: Dramatyczny spadek pewności we wszystkich modelach
- **Dim=64**: GAT wykazuje najniższą średnią pewność (0.2274)
- **GCN** utrzymuje względnie wysoką **pewność** nawet przy degradacji dokładności

## 5. Wnioski

### 5.1 Odporność na Destrukcję cech

1. **GAT**: Najodporniejszy na randomizację, zachowuje względnie wysoką dokładność
2. **GCN**: Średnia odporność, ale utrzymuje wysoką pewność predykcji
3. **GraphSAGE**: Najmniej odporny, największa degradacja wydajności

### 5.2 Problem Overfittingu

- Wszystkie modele wykazują silny **overfitting** na oryginalnym datasetcie
- **100% dokładność treningowa vs ~80% testowa** wskazuje na potrzebę regularyzacji
- Randomizacja cech pogłębia problem **overfittingu**

### 5.3 Mechanizmy Uwagi vs Agregacji

- **GAT (mechanizm uwagi)** wykazuje lepszą adaptację do zaburzonych cech
- **GCN (prosta agregacja)** utrzymuje stabilność pewności predykcji
- **GraphSAGE (sampling)** najbardziej zależny od prawidłowych cech sąsiedztwa

### 5.4 Implikacje dla Zastosowań Praktycznych

1. **Dla danych o stabilnych cechach**: Wszystkie modele są porównywalne
2. **Dla danych z zaburzonymi cechami**: GAT jest preferowany
3. **Gdy pewność predykcji jest kluczowa**: GCN może być lepszym wyborem
4. **Potrzeba lepszej regularyzacji** we wszystkich modelach