

This repository | Search

[Pull requests](#) [Issues](#) [Marketplace](#) [Gist](#) [chuanconggao](#) / [PrefixSpan-py](#)

Watch

1

Star

11

Fork

8

<> Code

Issues 0

Pull requests 0

Projects 0

Wiki

Insights

The shortest yet efficient implementation of famous frequent sequence mining algorithm PrefixSpan in Python 🐍.

9 commits

1 branch

0 releases

1 contributor

MIT

Branch: master

[New pull request](#)[Create new file](#)[Upload files](#)[Find file](#)[Clone or download](#)**chuanconggao** Add support for top-k algorithm

Latest commit e9dc8fa on 4 Mar

| | | |
|-------------------------------|---------------------------------|--------------|
| LICENSE | Create LICENSE | a year ago |
| README.md | Add support for top-k algorithm | 3 months ago |
| prefixspan.py | Add support for top-k algorithm | 3 months ago |

[README.md](#)

The shortest yet efficient PrefixSpan implementation in Python, in only 20 lines in core part.

Usage

Just replace the variable `db` with your own sequences, and variable `minsup` with your own minimum support threshold.

Features

Based on state-of-the-art [PrefixSpan](#) algorithm. Mining top-k patterns is also supported.

Tip

I strongly encourage using PyPy instead of CPython to run the script for best performance. In my own experience, it is 9x times faster in average.

