

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

<http://marjani.net/blog>

# داده های جریانی در یادگیری ماشین

<http://marjani.net>

<http://marjani.net/blog>



# فهرست مطالب

مقدمه

تاریخچه

نمونه ها و ویژگی های داده های جریانی

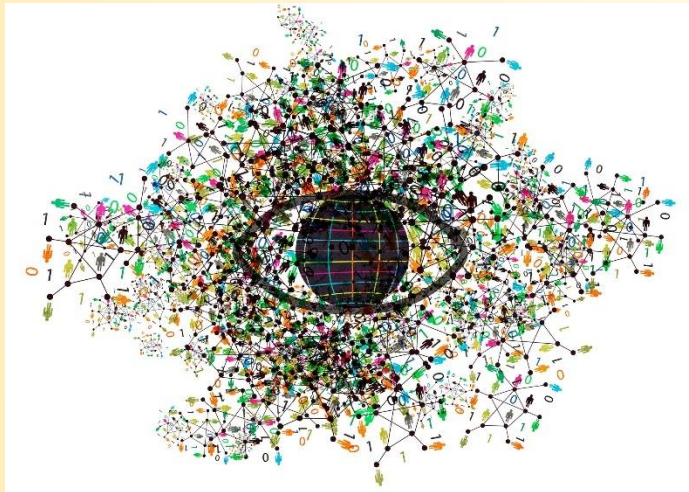
محدودیت های داده های جریانی

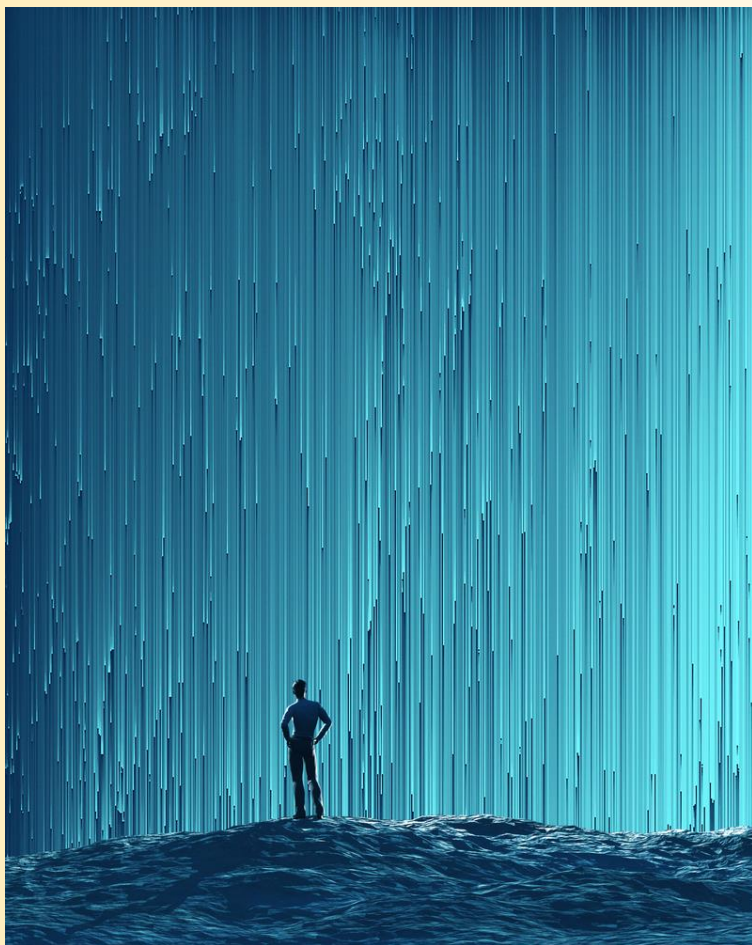
تعریف مفاهیم

رویکردهای محاسباتی در داده های جریانی

اعتبار سنجی در داده های جریانی

<http://marjani.net/blog>





## منظور از الگوریتم های داده های جریانی (جریان داده ها) چیست؟

**خلاصه:** الگوریتم هایی که با داده های پویا و نامحدود کار می کنند را الگوریتم های داده های جریانی می نامند.

**مفصل:** الگوریتمی است برای پردازش جریان داده ها که به صورت توالی از آیتم ها به عنوان ورودی هستند و از یک یا تعداد محدودی گذرگاه می توانند مورد بررسی قرار بگیرند. در واقع این الگوریتم دنباله ای از داده را به عنوان ورودی دریافت کرده و یک سری تابع روی آنها اعمال می کند. این الگوریتم حافظه مورد نیاز (کمتر از اندازه واقعی خود ورودی) و زمان پردازش هر آیتم را هم محدود می کند. این محدودیت ها به این معناست که این الگوریتم جواب تقریبی بر اساس یک خلاصه یا (طرحی) از جریان داده ها در حافظه تولید می کند.



- مطالعه الگوریتم جریان توسط مونرو و پاترسون در اوایل ۱۹۸۰ آغاز شد
- نایجل مارتین در ۱۹۸۲-۱۹۸۳ باعث شد، موضوعات مرتبط به الگوریتم‌های جریان برای اولین بار به رسمیت شناخته شود.
- در سال ۱۹۹۶ توسط نوگا آلون، یوسی ماتياس، و ماریو محبوب شد
- الگوریتم‌های نیمه جریانی در سال ۲۰۰۵ به عنوان یک فرمت از الگوریتم‌های جریانی ارائه شد



## نمونه های داده های جریانی



- جریان ترافیک در کلانشهرها و جاده ها
- هواشناسی و داده های اقلیمی
- داده های زیست محیطی
- شبکه های تلویزیونی و رادیویی
- شبکه های حسگر گوناگون
- علوم و فنون فضایی، جغرافیایی، نظامی
- مراقبت های ویژه پزشکی
- تجارت، بورس و بازار

# نمونه ها و ویژگی های داده های جریانی (فناوری اطلاعات)

## نمونه های دادهای جریانی (فناوری اطلاعات)

- انتقال داده ها در شبکه های رایانه ای (ترافیک شبکه)
- اطلاعات نشست (لاگ)
- امنیت سیستم ها (کنترل امنیت)
- کانال ها در شبکه های اجتماعی
- برنامه های اینترنتی و پادکست ها
- تراکنش های مالی / بانکی برخط
- انواع سیستم های مبتنی بر سنسور (مانند RFID)



# نمونه ها و ویژگی های داده های جریانی (فناوری اطلاعات)

## ویژگی های داده های جریانی

Real-time data

Continuous data

Ordered data

Large data

Non-ending data

Digital data

