



Classifying Self-Reported Rheumatoid Arthritis Flares Using Daily Patient-Generated Data From a Smartphone App: Exploratory Analysis Applying Machine Learning Approaches

استفاده از داده های روزانه تولیدشده توسط بیماران از طریق یک اپلیکیشن همراه برای طبقه بندی شدت تشدید آرتریت روماتوئیدی توسط بیمار: تجزیه و تحلیل کاوشی با استفاده از روش های یادگیری ماشین

مشخصات مقاله



Name:
JMIR
Formative
Research

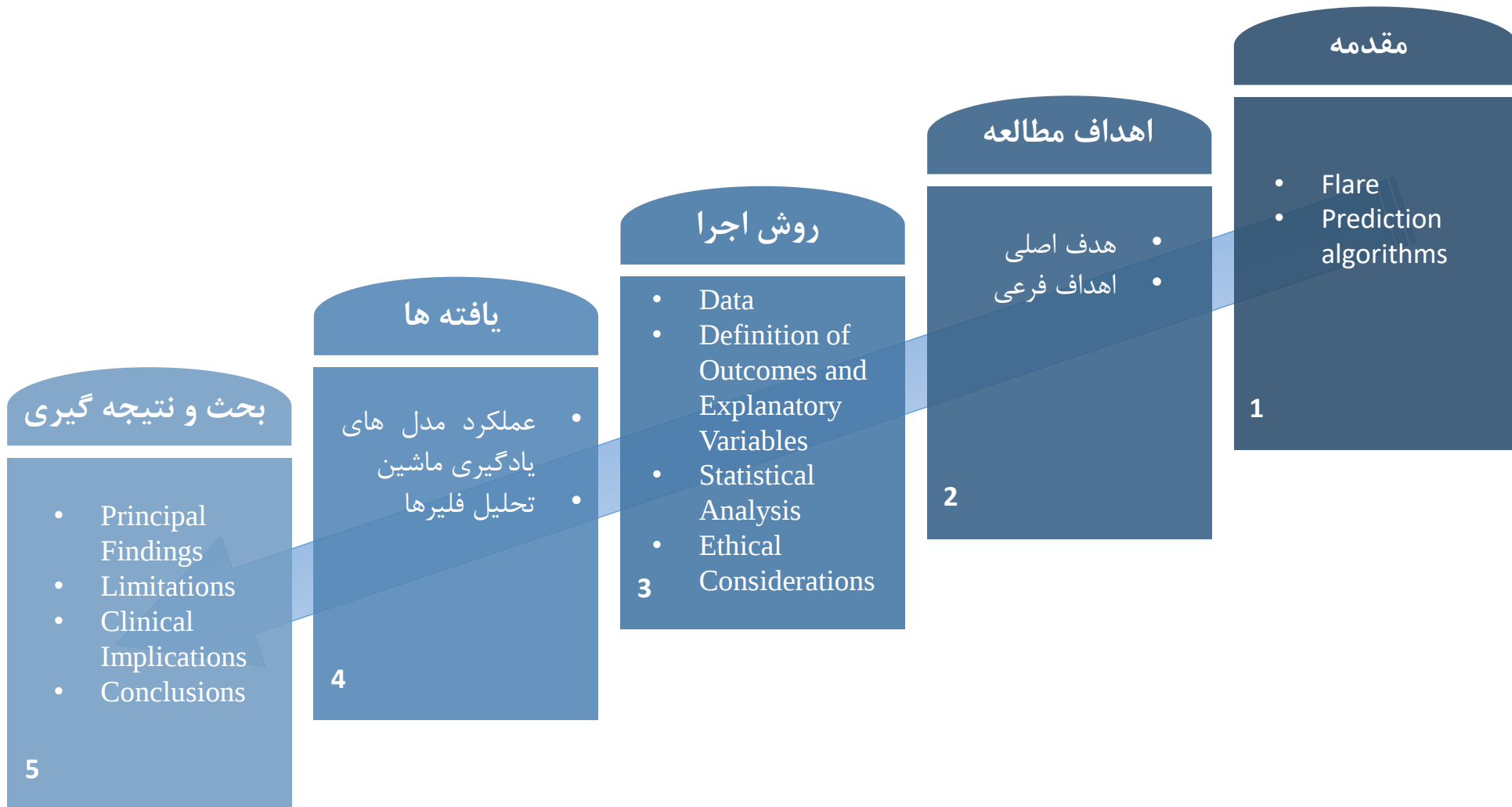
Category:
Q1

Indexing:
ESCI(ISI)
Scopus
PubMed

**Impact
Factor:**
2/000



فهرست مطالب





01

مقدمه

- شعله ور شدن
- الگوریتم های پیشبینی



Abbreviation	complete	meaning
AUC	area under the curve	مساحت زیرمنحنی
ML	machine learning	یادگیری ماشین
NPV	negative predictive value	ارزش پیشبینی منفی
PPV	positive predictive value	ارزش پیشبینی مثبت
RA	rheumatoid arthritis	آرتریت روماتوئید
REMORA	Remote Monitoring of Rheumatoid Arthritis	نظارت از راه دور آرتریت روماتوئید
ROC	receiver operating characteristics	ویژگی های عملیاتی گیرنده

بیماری آرتریت روماتوئید و اهمیت بررسی آن

مراجعه های پزشکی در هر 6 تا 12 ماه یکبار و مدیریت نشدن بسیاری از Flare ها
ضروری بودن شناسایی زودهنگام برای مداخلات درمانی

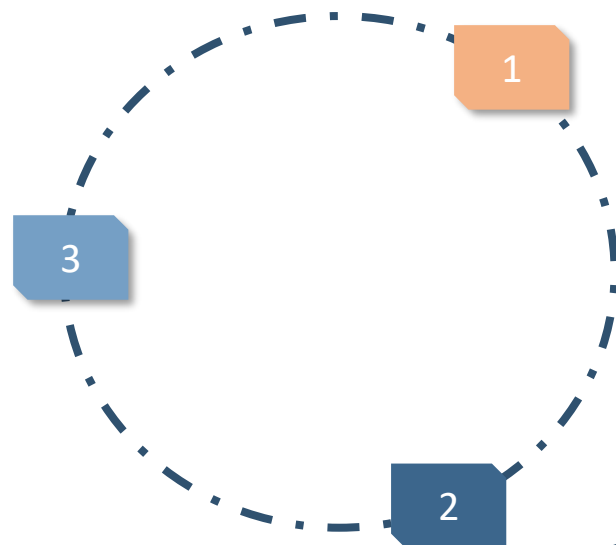
Flare ها باعث افزایش علائم ، کاهش عملکرد روزانه و آسیب ساختاری در بیماران

یک بیماری با شدت نوسانی و شامل شدن دوره هایی از تشدید بیماری (Flare)



نقش داده های تولیدشده توسط بیماران در مدیریت RA

ثبت علائم توسط بیماران از طریق گوشی های هوشمند و دستگاه های پوشیدنی و درک تغییرات بیماری توسط پزشکان



امکان جمع آوری روزانه یا هفتگی داده ها برای شناسایی زودهنگام Flare ها و پیشبینی آنها قبل از وقوع

امکان پذیر کردن مداخلات درمانی به موقع داده های بیمار محور مانند:

- مشاوره خودمدیریتی
- تنظیم داروها
- ویزیت آنلاین یا زودتر از موعد پزشک

استفاده از ML برای پیشبینی Flare ها



محدودیت داشتن روش های سنتی مانند رگرسیون لجستیک و شناسایی کردن الگوهای پیچیده تر توسط مدل های مدرن ML

وجود داشتن چالش هایی مانند عدم تعریف استاندارد Flare ها، تفاوت در روشهای جمع آوری داده و مدل های طبقه بندی



02

اهداف مطالعه

- هدف اصلی
- اهداف فرعی



هدف اصلی: بررسی امکان استفاده از یادگیری ماشین (ML) برای طبقه بندی Flare های گزارش شده توسط بیماران با استفاده از داده های یک اپلیکیشن موبایل

هدف فرعی: تحلیل تاثیر مقادیر آستانه مختلف بر دقت پیشبینی Flare ها

هدف فرعی: ارزیابی عملکرد ۳ مدل طبقه بندی دودویی و مقایسه آنها

هدف فرعی: ارائه مسیرهای پژوهشی آینده برای استفاده بهتر از داده های دیجیتال بیماران در پیشبینی Flare ها و بهبود مراقبت های پزشکی



روش اجرا

- ملاحظات اخلاقی
- **Data**
- Definition of Outcomes and Explanatory Variables
- Statistical Analysis
- Ethical Considerations



ملاحظات اخلاقی

1

بررسی و تاییدشدن ملاحظات اخلاقی توسط کمیته
اخلاق در پژوهش مرکزی منچستر

2

تکمیل کردن فرم های رضایت آگاهانه

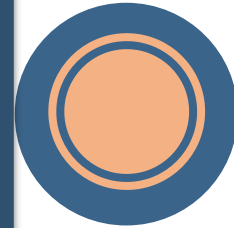


نوع مطالعه: تحلیل پساہنگی از داده های فاز اول مطالعه

REMORA

شرکت کنندگان: 20 بیمار مبتلا به آرتریت روماتوئید (RA)

مدت زمان: 3 ماه

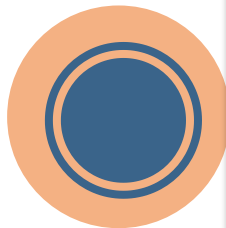


معیارهای ثبت شده: درد، عملکرد، خستگی، کیفیت خواب، سلامت

جسمی و روحی، توانایی مقابله

شرایط ورود به مطالعه: تشخیص RA توسط پزشک، تحت درمان در

کلینیک سرپایی، تمایل به مشارکت، ارائه رضایت نامه کتبی



Definition of Outcomes and Explanatory Variables

روش پردازش داده ها

- حذف شدن گزارش های روزانه که بعد از ۶ روز به فلیر متصل نبودند.
- درنظرگرفته شدن پاسخ های تکراری در یک روز به عنوان آخرین پاسخ
- ثبت شدن داده های علائم روی مقیاس 0 تا 10
- استفاده شدن همه علائم به عنوان متغیرهای پیوسته (continuous variables)

داده های ورودی مدل

- بررسی شدن داده های روزانه بیماران از ۷ روز قبل از هر گزارش فلیر.
- ویژگی های استخراج شده از علائم روزانه:
- حداقل، حداکثر، میانگین، انحراف معیار (SD) و شیب تغییرات (Slope)
- برای ۸ علامت مختلف (مانند درد و خستگی)

تعریف خروجی

- درنظر گرفتن هرگزارش هفتگی flare به عنوان یک متغیردودویی (Binary Outcome)
- بیماران تصمیم گرفتند که آیا دچار flare شده اند یا نه.

Definition of Outcomes and Explanatory Variables

چالش های مدل سازی

1

نبودن ابزارهای مناسب برای اعمال محدودیت
های یکنواختی (Monotonic Constraints)
در Python و R

2

بهبتر کردن تفسیر مدل و کاهش دقت

هدف تحلیل آماری

Statistical Analysis



استفاده از مدل های
یادگیری ماشین (ML)
برای پیشبینی وقوع Flare
ها



تشخیص و دسته بندی
هفته های flare و
غير flare بر اساس ویژگی
های علائم روزانه

Statistical Analysis



مقدمه

اهداف مطالعه

یافته ها

بحث و نتیجه گیری



مناسب بودن
روش بیز ساده
برای پیشبینی
های احتمالی



تشخیص دادن
الگوهای پیچیده
توسط مدل
جنگل تصادفی



قابل پیشبینی و
ساده بودن مدل
رگرسیون
لجستیک

Cutoffs	Psychometric properties			
	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Cutoff 1	0.50	0.90	0.65	0.83
Cutoff 2	0.60	0.80	0.53	0.85
Cutoff 3	0.70	0.74	0.49	0.87
Cutoff 4	0.80	0.72	0.51	0.90
Cutoff 5	0.90	0.43	0.37	0.92
Cutoff 6	0.88	0.50	0.39	0.92
Cutoff 7	0.87	0.60	0.44	0.92
Cutoff 8	0.83	0.70	0.51	0.92
Cutoff 9	0.60	0.80	0.53	0.85
Cutoff 10	0.50	0.90	0.65	0.83

حساسیت (Sensitivity): درصد فلیرهای واقعی که مدل به درستی شناسایی کرده است.

ویژگی (Specificity): درصد هفته های بدون فلیر که مدل به درستی شناسایی کرده است

ارزش پیشبینی مثبت (PPV): احتمال درست بودن پیشبینی های فلیر

ارزش های پیشبینی منفی (NPV): احتمال درست بودن پیشبینی های غیرفلیر





بررسی شدن مدل برای ۱۰ مقدار مختلف حساسیت و ویژگی
(بین ۰.۵ تا ۰.۹)

تحلیل
حساسیت
و ویژگی

مقایسه NPV و PPV برای بررسی میزان کاربرد
مدل در محیط های بالینی.

مقدمه

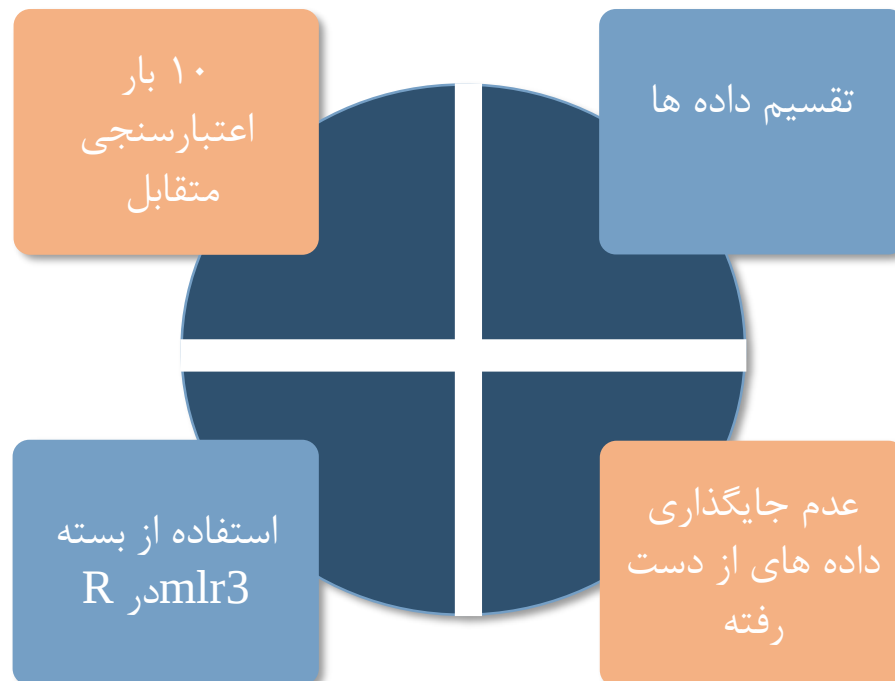
اهداف مطالعه

یافته ها

بحث و نتیجه گیری

Statistical Analysis

فرآیند آموزش:

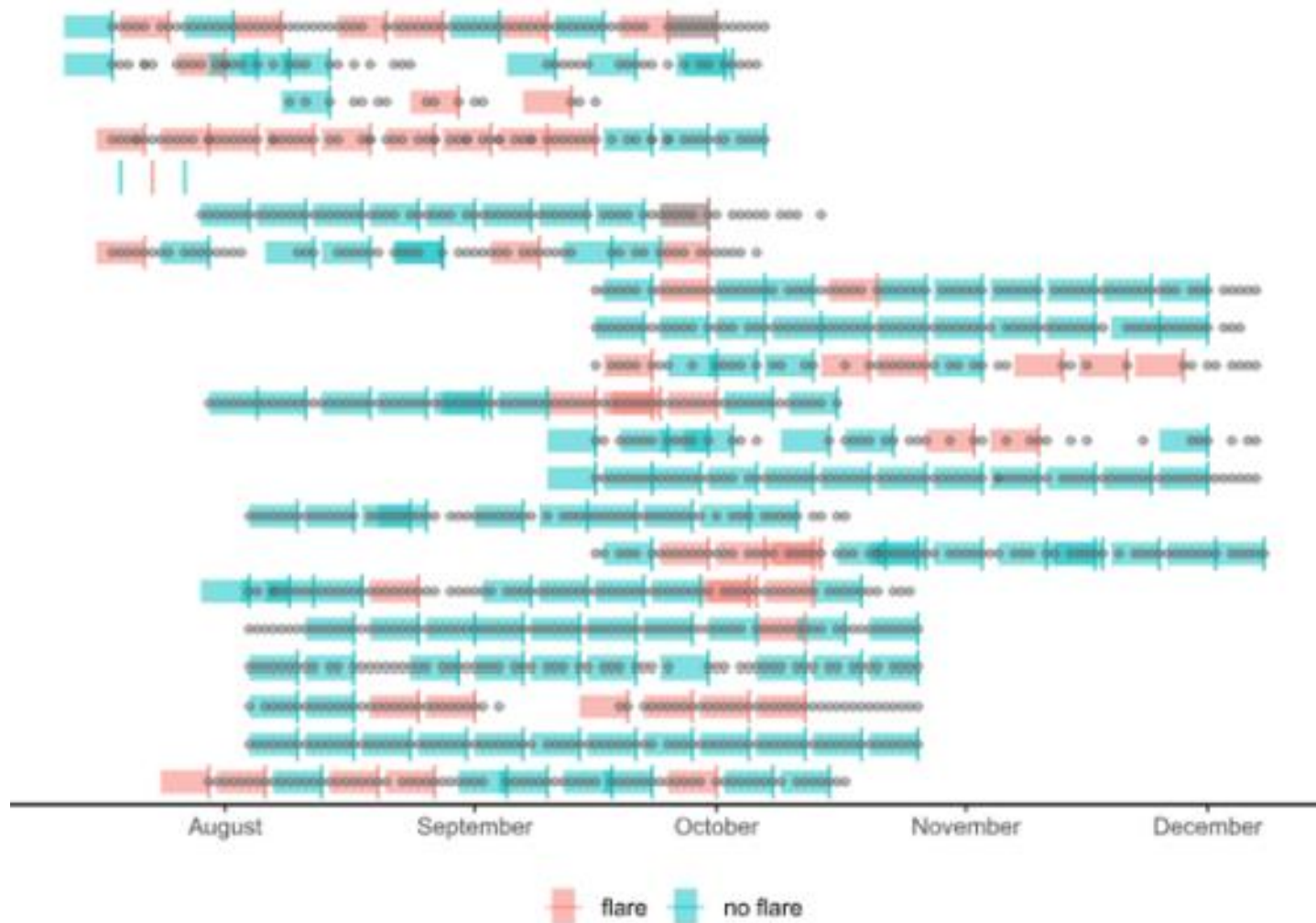




یافته ها

- عملکرد مدل های یادگیری ماشین
- تحلیل فلیرها





تحلیل Flare ها



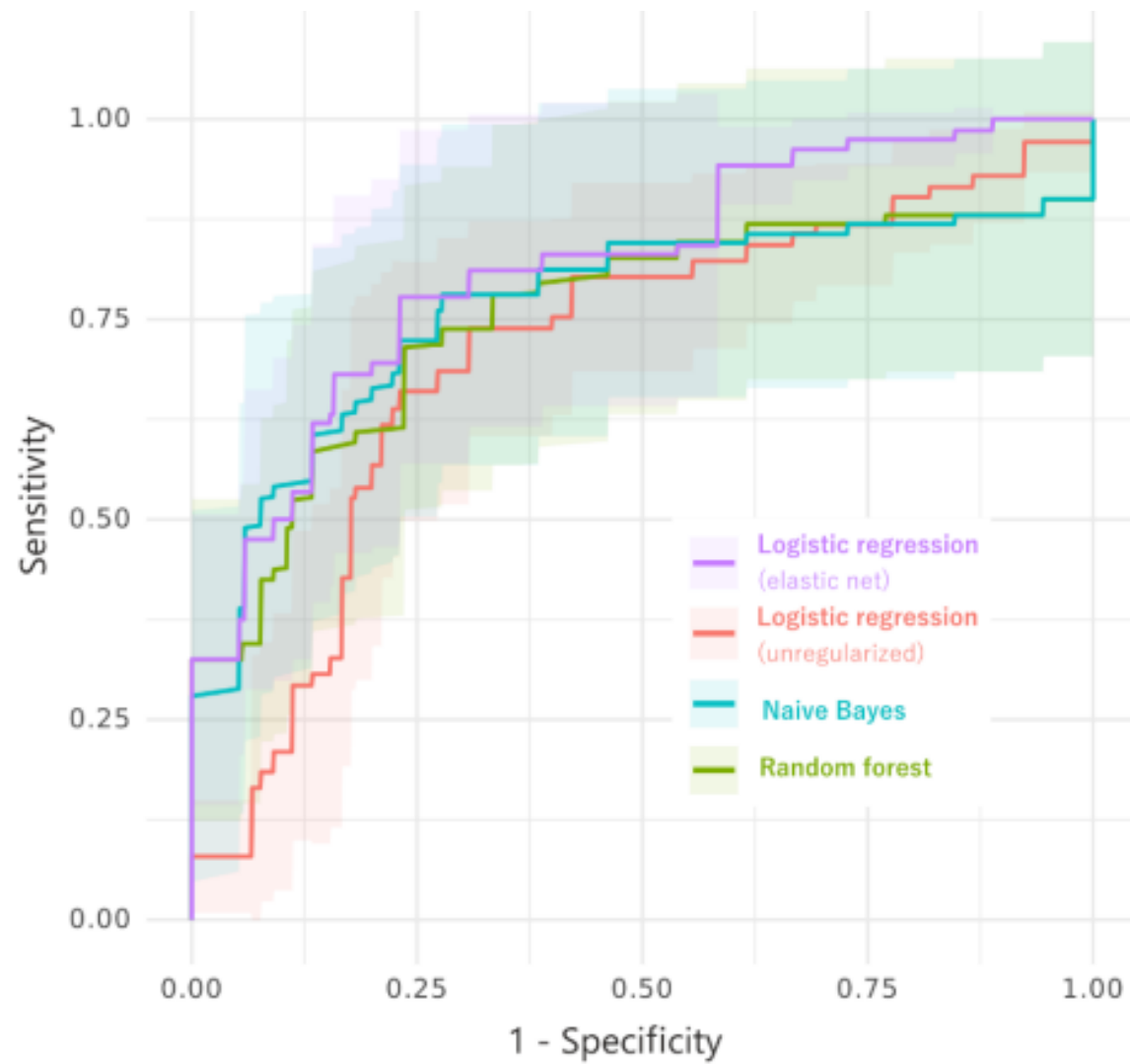
بیشترین تعداد Flare
گزارش شده توسط یک نفر:
9 مورد



تعداد کل Flare ها: 57
میانگین Flare ها: 2 مورد
(IQR: 0.75-4.25)



5 نفر هیچ Flare گزارش
نکردند.



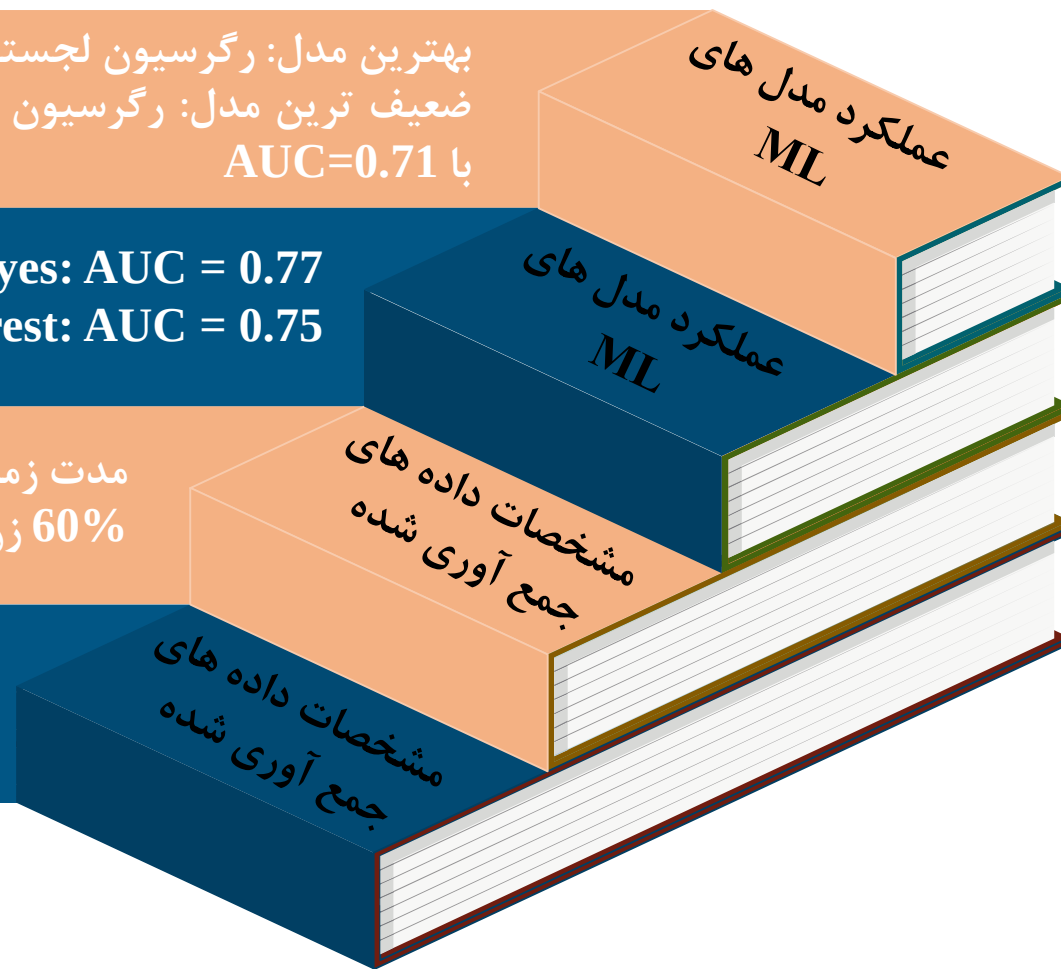
مشخصات داده های جمع آوری شده

بهترین مدل: رگرسیون لجستیک با $AUC=0.82$
 ضعیف ترین مدل: رگرسیون لجستیک بدون منظم سازی
 با $AUC=0.71$

Naïve Bayes: $AUC = 0.77$
 Random Forest: $AUC = 0.75$

مدت زمان پیگیری: 81 روز (IQR:79-82)
 60% زن (12 نفر)، میانگین سنی 57 سال

20 نفر شرکت کننده، 1325 گزارش
 روزانه، 213 گزارش هفتگی



Cutoffs	Psychometric properties			
	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Cutoff 1	0.50	0.90	0.65	0.83
Cutoff 2	0.60	0.80	0.53	0.85
Cutoff 3	0.70	0.74	0.49	0.87
Cutoff 4	0.80	0.72	0.51	0.90
Cutoff 5	0.90	0.43	0.37	0.92
Cutoff 6	0.88	0.50	0.39	0.92
Cutoff 7	0.87	0.60	0.44	0.92
Cutoff 8	0.83	0.70	0.51	0.92
Cutoff 9	0.60	0.80	0.53	0.85
Cutoff 10	0.50	0.90	0.65	0.83



01

بالاترین مقدار PPV: 0.65 ,NPV: 0.83

02

بالاترین مقدار PPV: 0.51 ,NPV: 0.92



بحث و نتیجه گیری

- Principal Findings
- Limitations
- Clinical Implications
- Conclusions
- دیدگاه من



Principal Findings



رگرسیون لجستیک با منظم سازی بهترین مدل
با $AUC = 0.82$ بود.

جنگل تصادفی به دلیل حجم کم داده ها
احتمالا دچار بیش برازش شد.

در بهترین آستانه، مدل توانست 3
مورد از هر 5Flare را به درستی
شناسایی کند.

NPV بالا به این معنی است که مدل در
تشخیص هفته های بدون Flare عملکرد
خوبی دارد.

Limitation

کاهش
تعمیم
قابلیت

حجم نمونه کوچک
و زمان پیگیری
محدود

متفاوت بودن نتایج
برای گروه های
دیگر

تنوع جمعیتی
پایین (اغلب
سفیدپوست بریتانیایی)

Clinical Implications



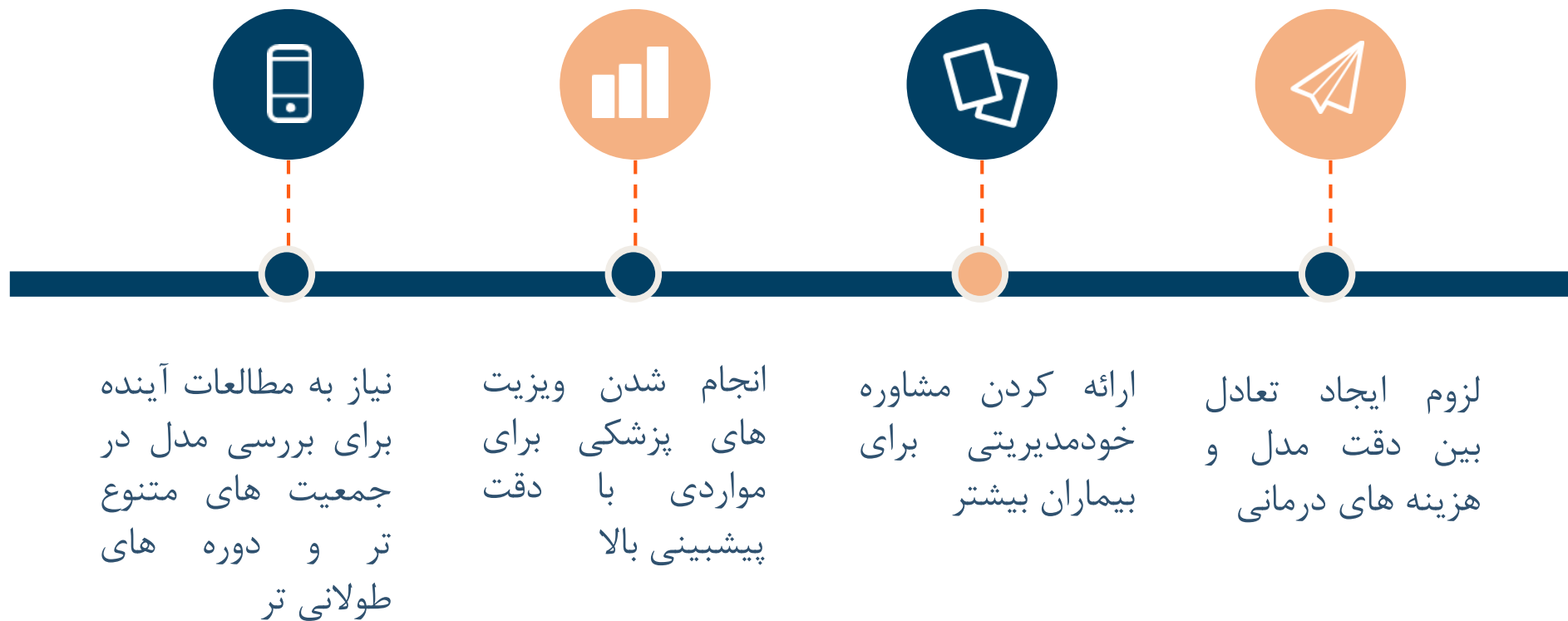
کاهش بار کاری
پزشکان از طریق
دخالت های خودکار



امکان
زودهنگام با
خودمدیریتی، مشاوره
تنظیم
دارو یا ویزیت پزشکی
زودتر از موعد



چالش ها و پیشنهادات



Conclusions

- امکان پذیر بودن پیشبینی فلیرهای آرتريت روماتوئيد با داده های روزانه بیماران و ML
- بهتر بودن عملکرد رگرسیون لجستیک منظم شده نسبت به سایر مدل ها
- نیازداشتن به داده های بیشتر و آزمایش های معتبر برای بهبود دقت پیشبینی
- ترکیب داده های سنسوری و گزارش های بیماران و درنهایت ارتقا دادن کیفیت پیشبینی



۲. بررسی مداخلات بالینی مبتنی بر پیشبینی ها



ارائه مشاوره آنلاین

پیشنهاد تنظیم دارو

ارسال هشدار به پزشک براساس
پیشبینی ها



با تشکر از توجه شما!
