

فاز اول پروژه ماشین لرنینگ

مرضیه کامکار
امیر سامان کریمی شرق
مهیار آلبالان

چکیده

در اینجا ما سعی داریم یک سیستم پیشنهاد دهنده با استفاده از تکنیک‌های ماشین لرنینگ و شبکه عصبی بسازیم و داده‌هایی که داریم از جنس دنبال کردن پادکست‌ها و کتگوری پادکست‌ها است. بخشی از فعالیتی که داریم با مقالات تطابق دارد و بخشی تطابق ندارد که در آینده به تفصیل آن‌ها می‌پردازیم.

مقاله‌ها

با توجه به داده‌ای که داریم و نوع این داده و محدودیت‌های محاسباتی‌ای که وجود دارد ما تصمیم گرفتیم به صورت استاتیک به داده‌ها نگاه کنیم و تمام مقاله‌ها از جنس *Cold – Collaborative filtering* است. با این وجود تفاوتی میان کار ما و کار این مقاله‌ها وجود دارد که در ادامه به آن‌ها خواهیم پرداخت.

داده‌ها

ما با استفاده از *API* اپلیکیشن کست باکس داده‌ی یوزرها و پادکست‌ها و پادکست‌ها را به دست آوردیم و یک شبکه دو بخشی که شامل پادکست‌ها و یوزرها است به ما داد. این شبکه یک این بخش در یک مقاله ذکر شده که گراف‌های جهت‌دار را چگونه می‌توان پیشنهاد داد و پیش‌بینی کرد

مدل سازی

اگر می‌خواستیم به شیوه مقاله‌ها پیش برویم باید ماتریس مجاورتی می‌ساختیم که در آن عددی با عنوان ریتینگ مطرح میشد ولی ما تصمیم گرفتیم که کلیت شبکه عصبی را بگیریم و به جای ریتینگ از کتگوری استفاده کنیم و میزان شباهت دو پادکست را با کتگوری و میزان تشابه شبکه یوزرها (همسایه اول) بسنجیم و با استفاده از یک ترشهلد در بالا و پایین این تشابه‌ها حاشیه امنی برای میزان تشابه پیدا کنیم. کلیت کار ما با مقالات تطابق دارد. هر دو شبکه‌های جهت‌داری داریم که به نوعی با استفاده از این شبکه‌ها و کمی اطلاعات متفاوت می‌خواهیم اطلاعات جدیدی اعم از کتگوری و اتصال را به صورت احتمالی پیش‌بینی کنیم

پیش پردازش

مقالاتی که خواندیم کار را به صورت نیمه کلاسیک انجام دادند و اطلاعات را به شکل ماتریس مجاورت داشتند ولی ما به شیوه جدیدی خواستیم در این مرحله نه تنها داده‌ها را تمیز کنیم بلکه یاد‌هایی به این گراف بر حسب شباهت اضافه کنیم که هندسه شبکه متفاوت شود و با مطالعه هندسه و زمینه کاری هر راس (کتگوری) به پردازش اطلاعات بپردازیم.