

## תיאור מילולי של האלגוריתם

נתחיל בשלב של אקספלורציה. בשלב זה נשתמש ב UCB על מנת להמליץ על תכנים, זאת בזמן שנוודא כי בהכרח נעמוד בדרישות של היצרנים בכל שלה האקספלורציה. בזמן שלב זה נבנה מטריצת תוחלות עבור כל המשתמשים ויצרני התוכן שכוללת את הרווח הממוצע עבור כל משתמש ותוכן שהומלץ לו.

בסוף שלב זה אנו אוספים את המידע שקיבלנו באקספלורציה. בעזרת מטריצת התוחלות, נבנה וקטור שיציג עבור כל משתמש את הבחירה האופטימלית עבורו (לפי היצרן עבורו יש את ערך התוחלת הכי גדול עבור כל משתמש). נבדוק את האפשרות של הסרת יצרני תוכן, כלומר יצרני תוכן שעדיף לנו לא לבחור בהן כלל ושיסומנו בתור יצרנים לא פעילים. על מנת לבדוק זאת נעבור על כל יצרני התוכן, עבור כל יצרן נניח כי אף משתמש לא יבחר בו, ונעדכן לפי זה את וקטור ההמלצות של המשתמשים. עבור הוקטור המקורי והוקטור החדש האלו נבדוק את תוחלת הרווח הממוצעת מסיבוב בודד (בעזרת המידע ממטריצת התוחלות והתפלגות בחירת המשתמשים) נשווה את שני התוחלות ביחס לאחוז הדרישות של היצרן באופן הבא: נגיד אנחנו בודקים אפשרות הסרה של יצרן מסוים שדורש לפחות 20 אחוזים מההמלצות בחלון הזמן. נבדוק האם:

$$(1+0.2) * x > (תוחלת בלי זרוע x)$$

אם התנאי מתקיים נבין שעדיף להסיר יצרן זה. במידה ובחרנו להסיר יצרן נעבור שוב מההתחלה על כל היצרנים משום שאולי כעת כדאי לנו להסיר יצרן נוסף.

כעת נשארנו עם היצרנים שאנחנו בהכרח רוצים לעמוד בדרישות שלהם. בשלב האקספלורציה המטרה שלנו היא שמתוך כל חלון זמן נבחר כמה שיותר פעמים ביצרן האופטימלי עבור כל משתמש, נעשה זאת באופן הבא: אם מספר הסיבובים שנותרו בסבב מסוים גדול יותר מסך הדרישות של כל יצרני התוכן שנותרו אז נבחר את היצרן האופטימלי. אחרת, אנחנו חייבים להמליץ אך ורק על היצרנים שלא עברו את הסף הדרוש של חלון הזמן הנוכחי, ועדיין ניקח את ההמלצות האופטימליות למשתמש (אשר לא כוללות בחירה של יצרנים שעברו את הסף) בעזרת מטריצת התוחלות שלנו. לדוגמא: נגיד יש לנו 3 יצרני תוכן, יצרן 0 דורש עוד 15 המלצות, יצרן 1 דורש עוד 7 המלצות ועבור יצרן 2 כבר סיפקנו את רף ההמלצות שהוא קבע, אם אנחנו בסיבוב ה-22 אז זה אומר שכעת אי אפשר לבחור ביצרן תוכן 2 עד סוף חלון הזמן (אפילו אם הוא האופטימלי עבור המשתמש) אלא חייבים להתחשב בדרישות (לפי ההנחה שלנו כי אנו רוצים לשמור על כל היצרנים שבחרנו). אם נעמוד בסף של יצרן נוסף אז נצטרך שוב לעדכן את הבחירות עבור כל משתמש. בדרך זו נבטיח כי נבחר כמה שיותר פעמים עבור כל משתמש את היצרן תוכן האופטימלי בשבילו ובכל זאת נעמוד בספים של כל יצרני התוכן שבחרנו אחרי האקספלורציה.