# **მსგავსი პროდუქტები**

**სატესტო დავალება მანქანური სწავლების ინჟინრის პოზიციაზე**

**პრობლემა**

ელექტრონული კომერციის სხვადასხვა პლატფორმებზე მომხმარებლებს ხშირად ექმნებათ პრობლემა სასურველი ნივთის შეძენისას. პრობლემა სხვადასხვა მიზეზებით შეიძლება იყოს გამოწვეული:

* მომხმარებელმა არ იცის ზუსტად რომელი პროდუქტი უნდა (იცის ჯგუფი, კატეგორია, მაგრამ ვერ აკეთებს საბოლოო არჩევანს)
* მომხმარებელი ვერ პოულობს სასურველ პროდუქტს
* მომხმარებლის სასურველი პროდუქტი აღარ არის მარაგში
* მომხმარებლის სასურველი პროდუქტის ფასი აღემატება მომხმარებლის ბიუჯეტს

და ა.შ.

**გამოსავალი**

ამ პრობლემის გადასაჭრელად და სამომხმარებლო გამოცდილების გასაუმჯობესებლად, საჭიროა ფუნქციონალი, რომელიც მომხმარებლებს შესთავაზებს მათი სასურველი პროდუქტის მსგავს პროდუქტებს.

ელექტრონული კომერციის არსებული პლატფორმების უმრავლესობა მსგავს პროდუქტებში აჩვენებს შემთხვევითობის პრინციპით შერჩეულ პროდუქტებს, იგივე კატეგორიიდან, რომელ კატეგორიაშიც არის მომხმარებლისთვის სასურველი პროდუქტი. ეს მიდგომა არაეფექტურია, ვინაიდან იგივე კატეგორიაში შესაძლოა იყოს რადიკალურად განსხვავებული მახასიათებლების მქონე პროდუქტები, მაგალითად უახლესი iPhone და ბიუჯეტური სმარტფონი ერთი და იგივე კატეგორიაშია, თუმცა ყიდვის პროცესში მათი მსგავს პროდუქტებად ჩათვლა მხოლოდ კატეგორიის გამო, არასწორია. შედეგად ვიღებთ მსგავსი პროდუქტების არაქმედით ფუნქციონალს, რომელიც ხშირ შემთხვევაში არარელევანტურ შედეგებს გვაძლევს.

მსგავსი პროდუქტების ეფექტური შეთავაზების პროცესი უნდა ითვალისწინებდეს არა მხოლოდ პროდუქტის კატეგორიას, არამედ პროდუქტის რიგ პარამეტრებს, რომლებიც განსაზღვრავს რეალურ მსგავსებას. მაგალითად ტელევიზორის შემთხვევაში ეს შეიძლება იყოს ეკრანის ზომა, გარჩევადობა, ინტერნეტთან კავშირის საშუალება, საფასო შუალედი და ა.შ.

**დავალება**

შენი დავალებაა შეიმუშავო სრულფასოვანი (end-to-end) API სერვისი, რომელსაც ექნება ერთი მეთოდი: *GetSimilarProducts (int ProductID, int count = 5)* რომელსაც გადაეცემა პროდუქტის უნიკალური იდენტიფიკატორი და დასაბრუნებელი პროდუქტების რაოდენობა (სტანდარტულად 5), ხოლო რეზულტატად დააბრუნებს count რაოდენობის მსგავს პროდუქტს, დალაგებულს მსგავსების ყველაზე მაღალი მაჩვენებლით დაბლისკენ. დაბრუნებული რეზულტატი უნდა მოიცავდეს როგორც პროდუქტის იდენტიფიკატორს, ისე მსგავსების ალბათობას.

ამოცანის შესრულებისას გაითვალისწინეთ, რომ მიზანი არ არის მსგავსების საუკეთესო ალგორითმის/მოდელის მიგნება, ამოცანის მიზანია სრული გადაწყვეტის შექმნა (მუშა API სერვისი), რომელიც იყენებს მსგავსი პროდუქტების დადგენის რომელიმე ალგორითმს/მოდელს და იძლევა საშუალებას ამ ალგორითმის/მოდელის სამომავლო გაუმჯობესებისათვის.

**მონაცემები**

ამოცანის გადასაწყვეტად საჭირო მონაცემები მოცემულია ML Test Assignment - Similar Products Data.xlsx ფაილში. ფაილში სულ 3 ჩანართია:

* **Categories -** ამ ჩანართზე მოცემულია პროდუქციის კატეგორიების სრული კატალოგი, ხის სტრუქტურით.
* **Products -** ამ ჩანართზე მოცემულია პროდუქციის სრული კატალოგი. თითოეულ ჩანაწერზე მოცემულია პროდუქტის უნიკალური იდენტიფიკატორი, კატეგორიის იდენტიფიკატორი, პროდუქტის სათაური და მოკლე აღწერა და ფასი
* **ProductSpecifications -** ამ ჩანართზე მოცემულია პროდუქტის მახასიათებლები. თითოეული პროდუქტისთვის მოცემულია ატრიბუტის იდენტიფიკატორი, ატრიბუტის დასახელება და ატრიბუტის მნიშვნელობა