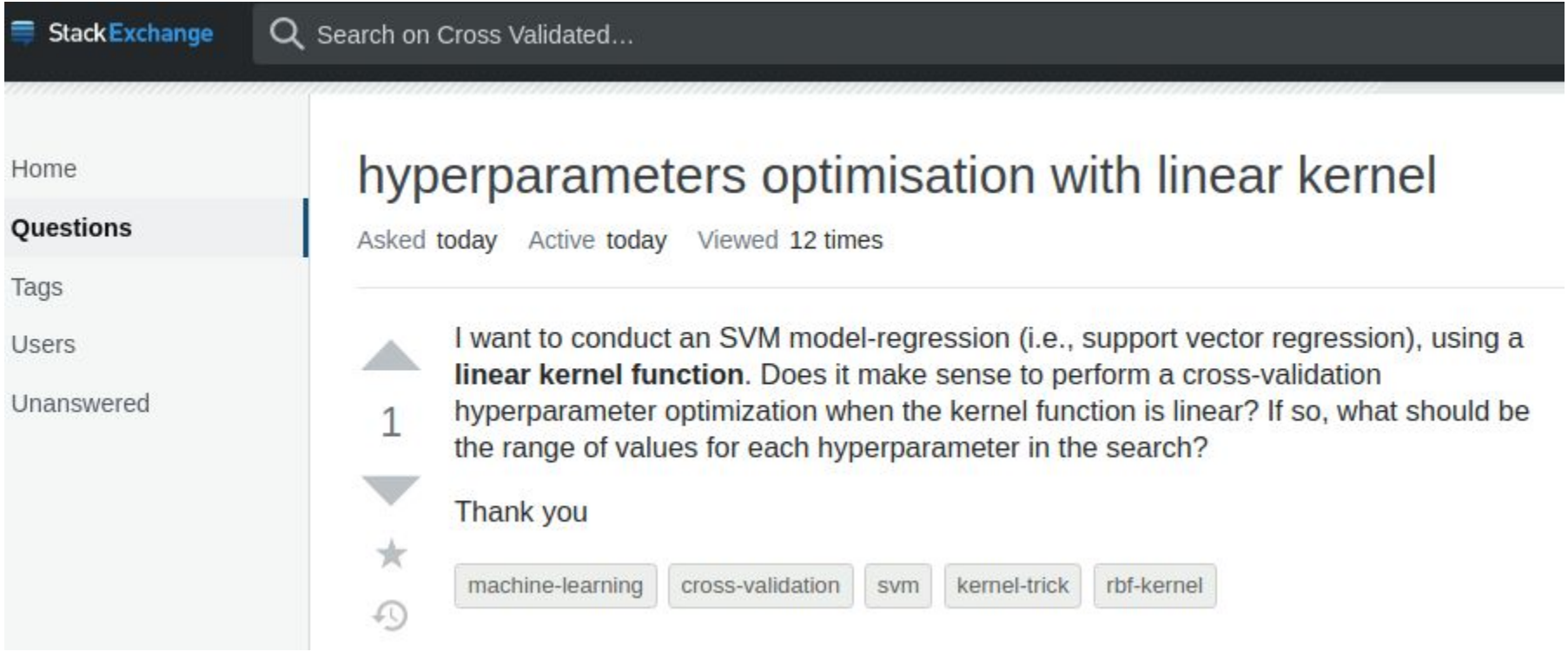
Auto-Tagging System

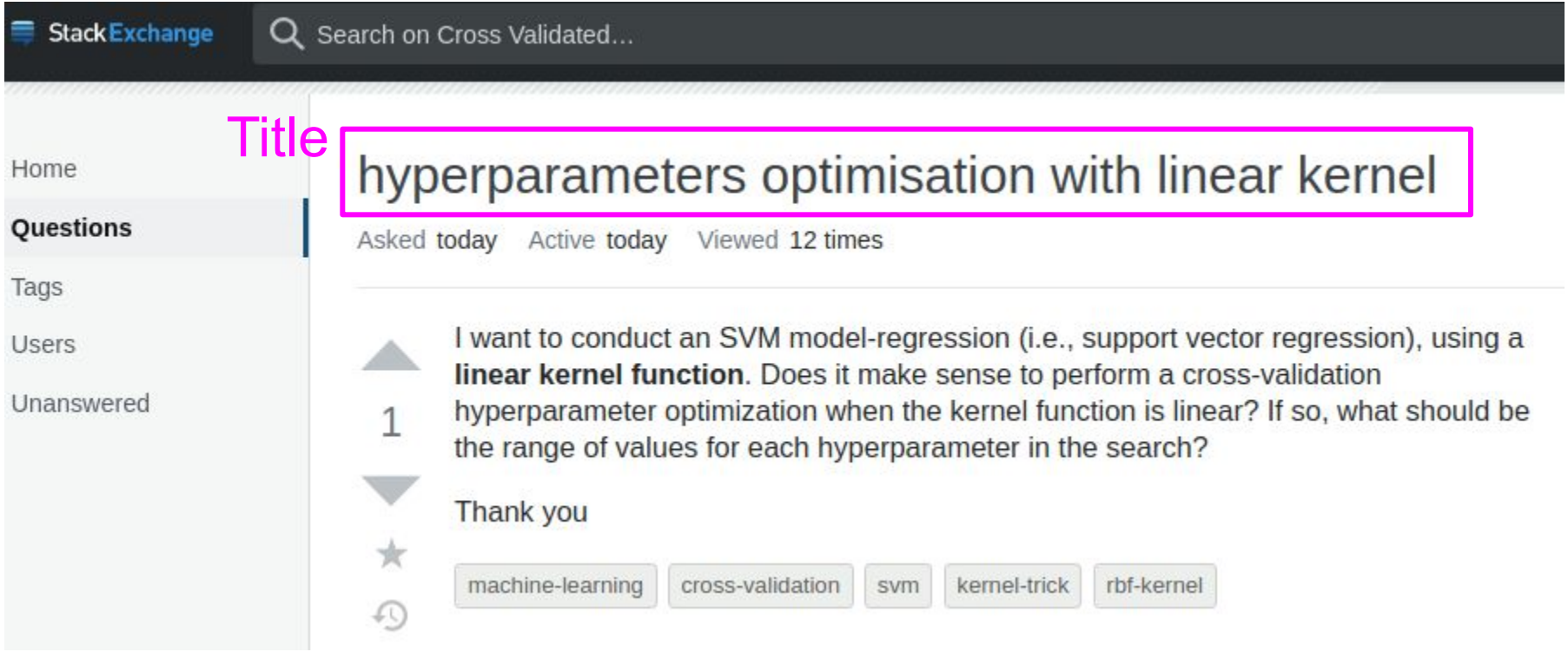
Auto-Tagging System

Understand the Business Problem



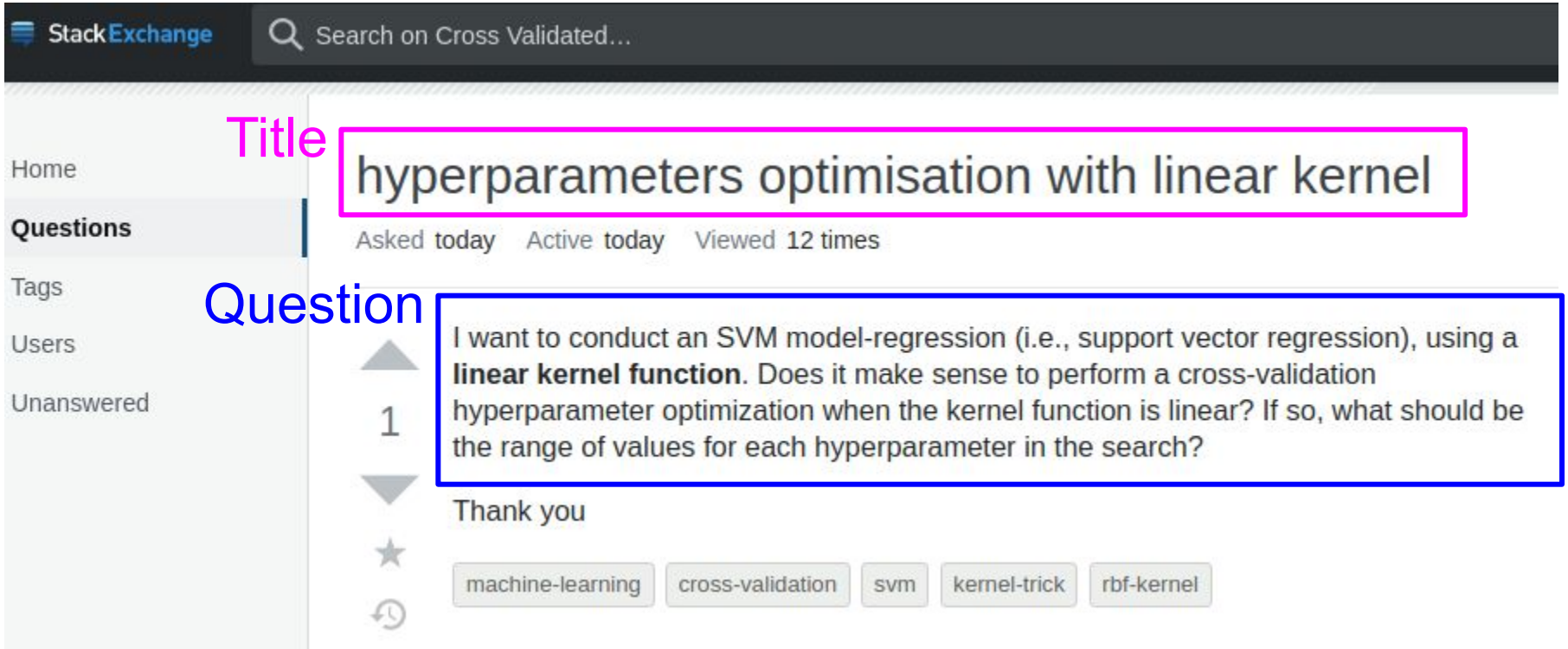
Analytics Vidhya
Learn everything about analytics

Understand the Business Problem



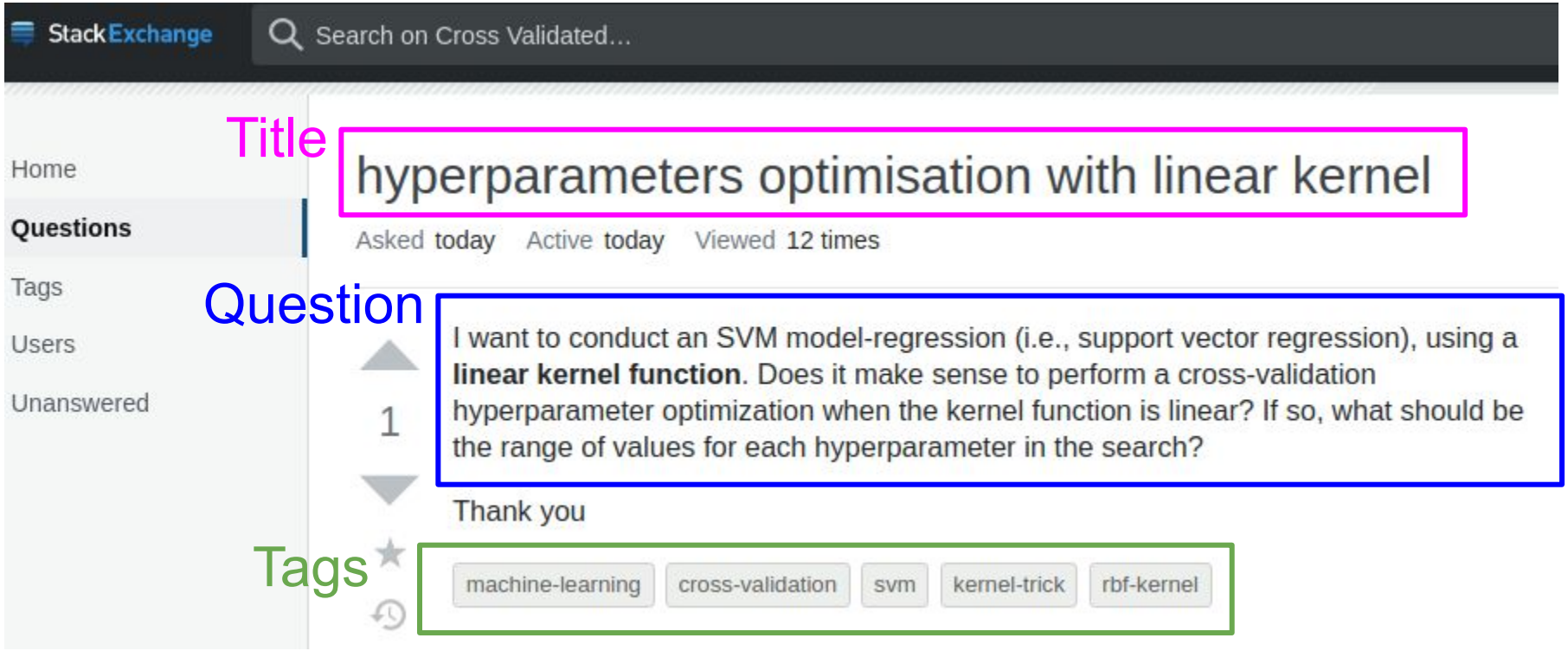
Analytics Vidhya
Learn everything about analytics

Understand the Business Problem



Analytics Vidhya
Learn everything about analytics

Understand the Business Problem



Analytics Vidhya
Learn everything about analytics

Understand the Business Problem

Business Problem: Untagged questions or questions with irrelevant tags lead to a large number of unanswered questions. The experts find it difficult to discover the right questions, and the users who post questions don’t get timely responses. This situation leads to a drop in the level of user-engagement on the platform

Data Science Problem

Data Science Problem

- Input: Text from the question

Data Science Problem

- Input: Text from the question
- Target Variable: Multiple Tags

Data Science Problem

- Input: Text from the question
- Target Variable: Multiple Tags
- Different from Binary and Multiclass Classification

“The rising environmental pollution screams out for a need to reduce fossil fuel consumption till a replacement for mobility doesn’t spread out in the world”

“The rising **environmental pollution** screams out for a need to reduce **fossil fuel** consumption till a replacement for mobility doesn’t spread out in the world”

Data Science Problem

- Input: Text from the question
- Target Variable: Multiple Tags
- Different from Binary and Multiclass Classification

X	**y**
X_1	t_1
X_2	t_2
X_3	t_3

Multiclass

Data Science Problem

- Input: Text from the question
- Target Variable: Multiple Tags
- Different from Binary and Multiclass Classification

X	**y**
X_1	t_1
X_2	t_2
X_3	t_3

Multiclass

X	**y**
X_1	$[t_2, t_5]$
X_2	$[t_1, t_3, t_2]$
X_3	$[t_4]$

Multilabel

Data Science Problem

- Input: Text from the question
- Target Variable: Multiple Tags
- Different from Binary and Multiclass Classification

X	y
X_1	t_1
X_2	t_2
X_3	t_3

Multiclass

X	y
X_1	$[t_2, t_5]$
X_2	$[t_1, t_3, t_2]$
X_3	$[t_4]$

Multilabel

t_1: 'python'

t_2: 'regression

t_3: 'numpy'

Thank You