**Recommendation System**

A recommendation system is a data filtering tool that recommends the most relevant items to a particular user or a customer using machine learning algorithms. In our case, items are colleges and users are high school students.

There are four main types of recommendation systems or engines: Content-Based filtering, Collaborative filtering, Knowledge-Based filtering and Hybrid recommendation systems. We will care the most about Knowledge-Based recommendation systems.

**Knowledge-Based Recommendation Systems (KB RS):**

A knowledge-based recommendation system is a system that make recommendations based on specific queries made by the user, not on a user’s rating history. It prompts the user to give some rules or guidelines on what the results should look like. The user may also give an example of an item. The system searches through its item database and tries to find a similar match. There are two main types of knowledge-based recommendation systems:

**1. Constraint-Based:**

It uses existing knowledge bases that contain explicit rules about how to relate user requirements with item features. It is very similar to an expert system in knowledge base systems. It is considered a constraint satisfaction problem.

**2. Case-Based:**

It depends on letting the user choose a target or an anchor item the algorithm finds a similar item to recommend. Results are usually treated as new target cases with some interactive modifications. This is similar to Content-Based systems where items similar to ones the user previously liked are suggested. The main difference is that most knowledge-based systems depend on the description of the items in the form of relational attributes in knowledge bases rather than as text keywords like in Content-Based systems.

**2.1. Similarity Metrics:**

Similarity metrics are needed to retrieve examples similar to the specified item. For continuous variables, similarity can be the difference between two numbers. For example, two different colleges have similar fees if the difference in fees is close to zero. It can also be more complex and require the use of statistical measures such as the standard deviation. When it comes to categorial variables, similarity is much more challenging. Domain hierarchies are often used in this case. Domain hierarchies are tree graphs where each category can have a parent category and subcategories. The similarity of two items can then be measured by the length of the path between the items.

**2.2. Critiquing:**

Once some result is found using similarity metrics and recommended to the user, the user is able to provide feedback and customizes the results to match what he is looking for. The user specifies a change request on some attributes of the item they like. The change request can be a directional critique or a replacement critique. A directional critique provides feedback on the direction in which the recommendation should be adjusted. For example, if the system recommended a college that is 100 km away, the user may request a shorter distance. On the other hand, a replacement critique suggests alternative items to replace the initially recommended item. For example, if the system recommended a computer engineering college, the user may suggest recommending a computer science college instead.

**KB RS Strengths:**

* Complex Item Domain: KB RS is very proficient when the items have many complex aspects to consider. Those aspects often need expert knowledge in the domain. KB RS effectively captures those complex aspects by utilizing knowledge bases crafted by domain experts. This strength is especially significant when the item is of great importance and requires a lot of thought, such as a university or a college.
* Avoids Cold Start Problem: User data is not needed because what the user wants is explicitly defined. The recommendation process is accurate and can start and work well without requiring existing rating data.

**KB RS Weaknesses:**

* Knowledge Acquisition Bottleneck: A big problem lies in the creation of the knowledge base. It requires the conversion of the knowledge possessed by domain experts into consumable representations.