

The background of the slide features a complex, abstract design. It includes a network of thin, light-colored lines forming a web-like structure. Scattered throughout are numerous small, colored dots in shades of green, blue, and orange. A prominent feature is a large, irregular polygonal shape in the center, outlined in a darker red or brown color. The overall color palette is muted, with a lot of grey and white space, accented by the colors of the data points and lines.

## **5.4. Mining Diverse Patterns: Mining Negative Correlations**

### **In-Lecture Question**

# Question

---

Consider the support-base and null-invariant definitions for negative patterns. Let  $\epsilon = 0.01$ . In a database containing  $10^9$  total transactions,  $10^6$  transactions contain item A and  $10^4$  transactions contain item B.  $10^2$  transactions contain both A and B. Which of the following is true?

- ☐ {A, B} is a negative pattern only by the support-based definition.
- ☐ {A, B} is a negative pattern only by the null-invariant definition.
- ☐ {A, B} is a negative pattern by both definitions.
- ☐ {A, B} is a negative pattern by neither definition.
- ☐ More information is needed to determine whether {A, B} is a negative pattern.

# Question

Consider the support-base and null-invariant definitions for negative patterns. Let  $\epsilon = 0.01$ . In a database containing  $10^9$  total transactions,  $10^6$  transactions contain item A and  $10^4$  transactions contain item B.  $10^2$  transactions contain both A and B. Which of the following is true?

- ☐ {A, B} is a negative pattern only by the support-based definition.
  - ☐ {A, B} is a negative pattern only by the null-invariant definition.
  - ☐ {A, B} is a negative pattern by both definitions.
  - ☐ {A, B} is a negative pattern by neither definition.
  - ☐ More information is needed to determine whether {A, B} is a negative pattern.
- 
- ☐ **Answer:** {A, B} is a negative pattern only by the null-invariant definition.
  - ☐ **Explanation:**  $\text{sup}(A) = 10^{-3}$ ,  $\text{sup}(B) = 10^{-5}$ ,  $\text{sup}(A \cup B) = 10^{-7} > \text{sup}(A) \times \text{sup}(B) = 10^{-8}$   
 $\Rightarrow$  {A, B} is not a negative pattern by the support-based definition.  
 $P(B|A) = P(A \cup B)/P(A) = 10^2/10^6 = 10^{-4}$ ,  $P(A|B) = P(A \cup B)/P(B) = 10^2/10^4 = 10^{-2}$   
 $(P(B|A) + P(A|B))/2 = (10^{-4} + 10^{-2})/2 = 0.00505 < 0.01$   
 $\Rightarrow$  {A, B} is a negative pattern by the null-invariant definition.