

Book Title	Encyclopedia of Machine Learning	
Book CopyRight - Year	2010	
Title	Frequent Itemset	
Author	Particle	
	Given Name	Hannu
	Family Name	Toivonen
	Suffix	
	Email	hannu.toivonen@cs.helsinki.fi
Affiliation	Division	Department of Computer Science
	Organization	University of Helsinki
	Street	P.O. Box 68 (Gustaf Hallstromin katu 2b)
	Postcode	FI-00014
	City	Helsinki
	Country	Finland

F

Frequent Itemset

HANNU TOIVONEN
Department of Computer Science,
University of Helsinki,
Helsinki, Finland

Synonyms

Frequent set

Definition

Frequent itemsets (Agrawal et al., 1993, 1996) are a form of [▶frequent pattern](#). Given examples that are sets of items and a minimum frequency, any set of items that occurs at least in the minimum number of examples is a frequent itemset.

For instance, customers of an on-line bookstore could be considered examples, each represented by the set of books he or she has purchased. A set of books, such as {"*Machine Learning*," "*The Elements of Statistical Learning*," "*Pattern Classification*,"} is a frequent itemset if it has been bought by sufficiently many customers. Given a frequency threshold, perhaps only 0.1 or 0.01% for an on-line store, *all* sets of books that have been bought by at least that many customers are called frequent. Discovery of all frequent itemsets is a typical data mining task. The original use has been as part of [▶association rule](#) discovery. [▶Apriori](#) is a classical algorithm for finding frequent itemsets.

The idea generalizes far beyond examples consisting of sets. The pattern class can be re-defined, e.g., to be (frequent) subsequences rather than itemsets; or original data can often be transformed to a suitable representation, e.g., by considering each discrete attribute-value pair or an interval of a continuous attribute as an individual item. In such more general settings, the term [▶frequent pattern](#) is often used. Another direction to generalize frequent itemsets is to consider other conditions than frequency on the patterns to be discovered; see [▶constraint-based mining](#) for more details.

Cross References

- ▶Apriori Algorithm
- ▶Association Rule
- ▶Constraint-Based Mining
- ▶Frequent Pattern

Recommended Reading

- Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD international conference on management of data, Washington, DC* (pp. 207–216). New York: ACM.
- Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., & Verkamo, A. I. (1996). Fast discovery of association rules. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), *Advances in knowledge discovery and data mining* (pp. 307–328). Menlo Park: AAAI Press.