



Jihočeská univerzita  
v Českých Budějovicích  
University of South Bohemia  
in České Budějovice

**University of South Bohemia**

**Faculty of Science**

Artificial Intelligence and Data Science, M.Sc.

## DESIGNING A BLOOM FILTER FOR A TARGET FALSE POSITIVE RATE

Submitted for the course **Information Theory**

**Professor:** Kaštovský Jan prof. RNDr. Ph.D.

Submitted by: María Isabel Sánchez-O'Mullony Martínez

Student ID: B24763

Date: November 24, 2025

Supervisor: Kaštovský Jan prof. RNDr. Ph.D.

## **Declaration**

I declare that I have written this report by myself and have only used the sources and aids mentioned, and that I have marked direct and indirect citations as such. This report has not been submitted prior for any other examination.

I agree that the results of this study work / report may be used free of charge for research and lecturing purposes.

# Contents

1	Introduction	4
	References	5

## List of Abbreviations

FPR False-Positive Rate

# 1 Introduction

A Bloom filter is a space-efficient probabilistic data structure that supports set membership queries, testing if an element is a member of a set[1]. The structure offers some interesting properties: It offers a compact probabilistic way to represent a set that can result in false positives (claiming an element to be part of the set when it is not), but never in false negatives (reporting an inserted element to be absent from the set). A Bloom filter of a fixed size can represent a set with an arbitrary large number of elements. Adding an element never fails, however, we need to take into account the increasing probability of false positive, and that deleting an element from filter is not possible[1, 2].

The objective of this project is to design and implement a Bloom filter that automatically chooses the optimal size ( $m$ ) and number of hash functions ( $k$ ) to achieve a target false-positive rate (FPR), given an expected number of elements  $n$ .

## References

- [1] S. Tarkoma, C. E. Rothenberg, and E. Lagerspetz, "Theory and practice of bloom filters for distributed systems," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 1, pp. 131–155, 2011.
- [2] GeeksforGeeks. (2025) Bloom filters - introduction and implementation. Last updated: 23 Jul 2025. [Online]. Available: <https://www.geeksforgeeks.org/python/bloom-filters-introduction-and-python-implementation/>