**A Minor Project Report on**

**Bias Detection in Hate Speech**

Report submitted to National Institute of Technology, Patna by  
Archana Kumari – 1906011  
Priyanshu Raj – 1906012

Garge Archana Atul – 1906124

Project done under the Supervision of  
**Dr. Jyoti Prakash Singh**  
Assistant Professor CSE Department, NIT Patna



**Department of Computer Science and Engineering**

**National Institute of Technology, Patna**

**Contents**

[Abstract 4](#_TOC_250008)

[Introduction 5](#_TOC_250007)

[Motivation 6](#_TOC_250006)

Related Work 7

Keywords and Definition 10

Methodologies 14

[Challenges faced 20](#_TOC_250005)

[Results 20](#_TOC_250004)

[Conclusion 21](#_TOC_250003)

[Future Scope 21](#_TOC_250002)

[References 22](#_TOC_250001)

[Appendix 24](#_TOC_250000)

**Abstract**

Gender Bias in hate speech relates to expressions which spread, incite, promote or justify hatred based on sex. Some groups of women are particularly targeted by sexist hate speech (notably young women, women in the media or women politicians), but every woman and girl is a potential target for online and offline sexist hate speech.  The increasing availability and use of Internet and social platforms have contributed to growing occurrences of sexist hate speech.In our method, we have tried to detect mitigate bias in hate speech. Our method provides 74% accuracy with deep learning models and 78% accuracy with gradient boosting.