



UNIVERSITY OF THE WITWATERSRAND

COMS3007

MACHINE LEARNING

---

# Assignment

---

*Authors:*

A.E Bank

N. Blundell

*Student Numbers:*

604124

1132328

May 20, 2018



# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Background</b>	<b>7</b>
2.1	Data Set . . . . .	7
2.2	About this design . . . . .	9



# Chapter 1

## Introduction

This is where we put an introduction. In addition to the assignment requirements, we explore and discuss methods of analysing categorical data.



# Chapter 2

## Background

### 2.1 Data Set

The data set used for this assignment is the Mushroom Data Set.[2] It is based on records from The Audubon Society Field Guide to North American Mushrooms. The samples are from 23 species of the Agaricus and Lepiota Family. Each sample has 22 attributes and is classified as either poisonous or edible. The attributes are as follows:

1. cap-shape	bell=b, conical=c, convex=x, flat=f, knobbed=k, sunken=s
2. cap-surface	fibrous=f, grooves=g, scaly=y, smooth=s
3. cap-colour	brown=n, buff=b, cinnamon=c, gray=g, green=r, pink=p, purple=u, red=e, white=w, yellow=y
4. bruises?	bruises=t, no=f
5. odor	almond=a, anise=l, creosote=c, fishy=y, foul=f, musty=m, none=n, pungent=p, spicy=s
6. gill-attachment	attached=a, descending=d, free=f, notched=n
7. gill-spacing	close=c, crowded=w, distant=d
8. gill-size	broad=b, narrow=n
9. gill-colour	black=k, brown=n, buff=b, chocolate=h, gray=g, green=r, orange=o, pink=p, purple=u, red=e, white=w, yellow=y
10. stalk-shape	enlarging=e, tapering=t
11. stalk-root	bulbous=b, club=c, cup=u, equal=e, rhizomorphs=z, rooted=r, missing=?
12. stalk-surface-above-ring	fibrous=f, scaly=y, silky=k, smooth=s
13. stalk-surface-below-ring	fibrous=f, scaly=y, silky=k, smooth=s
14. stalk-colour-above-ring	brown=n, buff=b, cinnamon=c, gray=g, orange=o, pink=p, red=e, white=w, yellow=y
15. stalk-colour-below-ring	brown=n, buff=b, cinnamon=c, gray=g, orange=o, pink=p, red=e, white=w, yellow=y
16. veil-type	partial=p, universal=u
17. veil-colour	brown=n, orange=o, white=w, yellow=y
18. ring-number	none=n, one=o, two=t
19. ring-type	cobwebby=c, evanescent=e, flaring=f, large=l, none=n, pendant=p, sheathing=s, z one=z
20. spore-print-colour	black=k, brown=n, buff=b, chocolate=h, green=r, orange=o, purple=u, white=w, yellow=y
21. population	abundant=a, clustered=c, numerous=n, scattered=s, several=v, solitary=y
22. habitat	grasses=g, leaves=l, meadows=m, paths=p, urban=u, waste=w, woods=d

The Field Guide apparently (it is not free) states that there is no simple rule for determining whether a mushroom from one of these families is edible or not.



## 2.2 About this design

This is a simple report template with the UCT logo. Feel free to use/modify it to suit your needs. Variables that need to be altered have been commented to make modifications easier. For example if you need to change the university logo, look for the comment `% University Logo` in this file and then make appropriate modifications in that line.

A Table of Contents and a bibliography have also been implemented. To add entries to your bibliography, simply edit `biblist.bib` in the root folder and then use the `\cite{...}` command in `main.tex` [1]. The Table of Contents will be updated automatically.

I hope that you find this template both visually appealing and useful.

— Lin



# Bibliography

- [1] Help on BibTeX entry types. <http://nwalsh.com/tex/texhelp/bibtex-7.html>. Accessed: 2015-03-12.
- [2] The Audubon Society Field Guide to North American Mushrooms  
Jeff Schlimmer.  
. Mushroom data set  
, the audubon society field guide to north american mushrooms (1981). g. h. lincoff (pres.), new york: Alfred a. knopf. <https://archive.ics.uci.edu/ml/datasets/mushroom>. Accessed: 2018-04-11.