Text Classification (Newspaper) using Naive Bayes

Sheryl Mathew (11627236)

26 October, 2018

## Data Collection

library(jsonlite)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(kableExtra)  
library(formatR)  
  
api = "https://content.guardianapis.com/search?"  
api\_key = "0e9a8c4e-1866-4503-a9eb-70e391c499c4"  
page\_size = "200"  
  
# pages = c(1:5) queries =  
# c('business','sports','entertainment','economy','politics','science','crime','health','art','technology')  
  
pages = c(1:2)  
queries = c("business", "sports")  
  
news\_data = data.frame()  
data\_count = list()  
sum\_data\_count = list()  
  
for (query in queries) {  
 for (page in pages) {  
 url = paste(api, "q=", query, "&page-size=",   
 page\_size, "&page=", page, "&show-fields=body",   
 "&api-key=", api\_key, sep = "")  
 json = fromJSON(url)  
 body = as.data.frame(json$response$results$fields)  
 data = as.data.frame(json$response$results)  
 data = subset(data, select = -c(fields))  
 data = cbind(data, body)  
 data\_count = append(data\_count, nrow(data))  
 news\_data = rbind(news\_data, data)  
 }  
   
 # select\_data = select(data,  
 # c('sectionName','body'))  
 # print(kable(head(select\_data), format =  
 # 'latex', booktabs = T,  
 # caption=paste('Table  
 # containing',toupper(query),' News'))  
 # %>% kable\_styling(latex\_options =  
 # c('striped','hold\_position')))  
 # cat('\n')  
   
 sum\_data\_count = append(sum\_data\_count,   
 sum(as.integer(data\_count)))  
 data\_count = 0  
}

print\_data = function(query, sum\_data\_count) {  
 print(paste("Number of rows in ", toupper(query),   
 " queries: ", sum\_data\_count))  
 cat("\n")  
}  
mapply(print\_data, query = queries, sum\_data\_count = sum\_data\_count)

## [1] "Number of rows in BUSINESS queries: 400"  
##   
## [1] "Number of rows in SPORTS queries: 400"

## $business  
## NULL  
##   
## $sports  
## NULL

cat("\n")

print(paste("Total Number of rows in News: ",   
 nrow(news\_data)))

## [1] "Total Number of rows in News: 800"

cat("\n")

summary(news\_data)

## id type sectionId   
## Length:800 Length:800 Length:800   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
## sectionName webPublicationDate webTitle   
## Length:800 Length:800 Length:800   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
## webUrl apiUrl isHosted pillarId   
## Length:800 Length:800 Mode :logical Length:800   
## Class :character Class :character FALSE:800 Class :character   
## Mode :character Mode :character Mode :character   
## pillarName body   
## Length:800 Length:800   
## Class :character Class :character   
## Mode :character Mode :character

## Data Cleaning

library(tidyverse)

## -- Attaching packages ------------------------------------------------------------------------------------------ tidyverse 1.2.1 --

## v ggplot2 3.0.0 v readr 1.1.1  
## v tibble 1.4.2 v purrr 0.2.5  
## v tidyr 0.8.1 v stringr 1.3.1  
## v ggplot2 3.0.0 v forcats 0.3.0

## -- Conflicts --------------------------------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x purrr::flatten() masks jsonlite::flatten()  
## x dplyr::lag() masks stats::lag()

library(reshape)

##   
## Attaching package: 'reshape'

## The following objects are masked from 'package:tidyr':  
##   
## expand, smiths

## The following object is masked from 'package:dplyr':  
##   
## rename

library(qdap)

## Loading required package: qdapDictionaries

## Loading required package: qdapRegex

##   
## Attaching package: 'qdapRegex'

## The following object is masked from 'package:ggplot2':  
##   
## %+%

## The following object is masked from 'package:dplyr':  
##   
## explain

## The following object is masked from 'package:jsonlite':  
##   
## validate

## Loading required package: qdapTools

##   
## Attaching package: 'qdapTools'

## The following object is masked from 'package:dplyr':  
##   
## id

## Loading required package: RColorBrewer

##   
## Attaching package: 'qdap'

## The following object is masked from 'package:reshape':  
##   
## condense

## The following object is masked from 'package:forcats':  
##   
## %>%

## The following object is masked from 'package:stringr':  
##   
## %>%

## The following object is masked from 'package:purrr':  
##   
## %>%

## The following object is masked from 'package:tidyr':  
##   
## %>%

## The following object is masked from 'package:kableExtra':  
##   
## %>%

## The following object is masked from 'package:dplyr':  
##   
## %>%

## The following object is masked from 'package:base':  
##   
## Filter

library(textclean)

##   
## Attaching package: 'textclean'

## The following objects are masked from 'package:qdap':  
##   
## check\_text, mgsub, replace\_contraction, replace\_number,  
## replace\_ordinal, replace\_symbol, strip, sub\_holder

news = news\_data %>% select(c("id", "pillarName",   
 "sectionName", "webTitle", "body")) %>%   
 rename(c(pillarName = "type", sectionName = "section",   
 webTitle = "title")) %>% na.omit()

abv = c("bn")  
rep = c("billion")  
  
clean\_function = function(data) {  
 data = str\_replace\_all(data, "<.\*?>",   
 "")  
 data = lapply(data, function(x) gsub("[[:punct:]]",   
 " ", x))  
 data = str\_replace\_all(data, " – ", " ")  
 data = gsub("\"", " ", data)  
 data = replace\_curly\_quote(data)  
 data = replace\_symbol(data)  
 data = replace\_money(data)  
 data = replace\_abbreviation(data)  
 data = replace\_word\_elongation(data)  
 data = replace\_contraction(data)  
 data = replace\_incomplete(data)  
 data = replace\_date(data)  
 data = replace\_ordinal(data)  
 data = replace\_number(data)  
 data = iconv(data, from = "UTF-8", to = "ASCII//TRANSLIT")  
 data = tolower(data)  
 data = trimws(data)  
 data = str\_squish(data)  
 data = str\_replace\_all(data, "[^a-zA-Z\\s]",   
 "")  
}  
  
news$body = clean\_function(news$body)

## Warning in NextMethod(.Generic): probable complete loss of accuracy in  
## modulus

## Warning in split\_digits(x): probable complete loss of accuracy in modulus  
  
## Warning in split\_digits(x): probable complete loss of accuracy in modulus  
  
## Warning in split\_digits(x): probable complete loss of accuracy in modulus

print("Columns in the News data before cleaning:")

## [1] "Columns in the News data before cleaning:"

names(news\_data)

## [1] "id" "type" "sectionId"   
## [4] "sectionName" "webPublicationDate" "webTitle"   
## [7] "webUrl" "apiUrl" "isHosted"   
## [10] "pillarId" "pillarName" "body"

unclean\_data = select(news\_data, c("id",   
 "pillarName", "sectionName", "webTitle",   
 "body"))  
  
kable(head(unclean\_data), format = "latex",   
 booktabs = T, caption = "Table containing details of Auto CSV") %>%   
 kable\_styling(latex\_options = c("striped",   
 "hold\_position", "scale\_down"))

print("Columns in the News data after cleaning:")

## [1] "Columns in the News data after cleaning:"

names(news)

## [1] "id" "type" "section" "title" "body"

clean\_data = select(news, c("type", "section",   
 "body"))  
  
# kable(head(clean\_data), format =  
# 'latex', booktabs = T, caption='Table  
# containing details of Auto CSV') %>%  
# kable\_styling(latex\_options =  
# c('striped','hold\_position','scale\_down'))

print("Sample of Unclean News Article Body:")

## [1] "Sample of Unclean News Article Body:"

cat("\n")

news\_data$body[2]

## [1] "<p>Niamh Black left school at 16 and spent four years working in a cafe before applying to become an apprentice dental nurse at Dental22 in Retford, Nottinghamshire.</p> <p>She is now a year into a course, and will be qualified within another year. She explains that she took the leap from working in hospitality to training as an apprentice because it offered far better career prospects.</p> <p>“I didn’t know what I wanted to do when I left school, so I just started working at a local cafe,” she recalls.</p> <p>“It was a good job, don’t get me wrong, but I couldn’t really see it going anywhere. I wanted to get a proper qualification that would lead to a decent career.”</p> <p>For dentist Margaret Naylor, Black’s experience sums up how apprenticeships empower people of any age to train for a new career. She believes they open up a wider supply of cost-effective talent to employers. It is a point that the <a href=\"https://www.theguardian.com/invest-in-an-apprentice/2018/oct/26/everything-employers-need-to-know-about-apprenticeships\">National Apprenticeship Service</a>, which supports England’s apprenticeship delivery, is keen to make clear to employers: apprenticeships are not limited to school leavers.</p> <figure class=\"element element-image element--supporting\" data-media-id=\"53367990a844b511313e6a08113fcfb1f9d1f754\"> <img src=\"https://media.guim.co.uk/53367990a844b511313e6a08113fcfb1f9d1f754/0\_0\_5600\_3360/1000.jpg\" alt=\"Dentist putting on surgical gloves, preparing to use dentistry equipmentGettyImages-951629888\" width=\"1000\" height=\"600\" class=\"gu-image\" /> <figcaption> <span class=\"element-image\_\_caption\">Apprentices can find themselves professionally qualified in the time it takes to complete an undergraduate course.</span> <span class=\"element-image\_\_credit\">Photograph: Eugenio Marongiu/Getty Images</span> </figcaption> </figure> <p>In fact, there is a special provision to keep the cost of recruitment low, so an SME, such as Dental22, which is looking to balance its wage bill, is not required to match the salary an older person may have been earning in another position. Instead, a first-year apprentice is paid the minimum wage of £3.70 per hour, regardless of age, minimising the burden on the employer.</p> <p>For Naylor, though, the main benefit apprentices offer is flexibility in training people to her high standards, rather than hoping that qualified staff have already been trained well at another practice.</p> <p>“Apprenticeships are all about flexibility. They allow people to change jobs and qualify in another area, just as much as they help school leavers get professional qualifications while they earn,” she says.</p> <p>“For us, as a business, the beauty of apprenticeships is that we can mould employees to meet our high standards. Sometimes we find that if we hire dental nurses who have qualified elsewhere, they’re not quite up to our standards.</p> <p>“It means we can end up having to retrain people, so for us, it just makes a lot more sense to train our own apprentices.”</p> <aside class=\"element element-rich-link element--thumbnail\"> <p> <span>Related: </span><a href=\"https://www.theguardian.com/invest-in-an-apprentice/2018/oct/26/many-of-the-partners-at-our-firm-started-as-apprentices\">‘Many of the partners at our firm started as apprentices’</a> </p> </aside> <p>Andi Martin is an example of a newly qualified dental nurse who has been trained to a high level within Dental22 over the past four years. She finished a level 3 NVQ in dental nursing in January, having joined straight after finishing her A-levels.</p> <p>It was the cost of university, and the size of the loan she would need to pay, that convinced Martin to become an apprentice.</p> <p>“I looked in to the cost of university and it would have been in the £30,000 zone,” she says.</p> <p>“That’s a scary amount of money when you don’t know for sure whether there will be a job at the end of it. So I went for the option of earning while I was learning – where the training is paid for you and you’re not left with a pile of debt.”</p> <p>Taking the apprenticeship route meant she would be professionally qualified by the time she would have been graduating. She’d also be without debt and already employed.</p> <p>In Naylor’s experience, training up an apprentice, such as Martin, not only benefits her business but also the wider dentistry and healthcare industry, because many apprentices carry on their studies, meaning the sector has a constant flow of higher-qualified staff.</p> <p>“A lot of apprentices qualify as dental nurses and then go on to study to become dental therapists,” she says.</p> <p>“A therapist can do a lot more work than a nurse: they can do restorations, fillings, take impressions and remove children’s teeth. Sometimes, dental nurses go on to become therapists or take other roles within the NHS.</p> <p>“By becoming apprentices and getting qualified, young people get access to these new career paths, and the industry gets a steady flow of highly qualified staff.”</p>"

strwrap(news\_data$body[2], width = 16)

## [1] "<p>Niamh Black"   
## [2] "left school at"   
## [3] "16 and spent"   
## [4] "four years"   
## [5] "working in a"   
## [6] "cafe before"   
## [7] "applying to"   
## [8] "become an"   
## [9] "apprentice"   
## [10] "dental nurse at"   
## [11] "Dental22 in"   
## [12] "Retford,"   
## [13] "Nottinghamshire.</p>"   
## [14] "<p>She is now a"   
## [15] "year into a"   
## [16] "course, and"   
## [17] "will be"   
## [18] "qualified"   
## [19] "within another"   
## [20] "year. She"   
## [21] "explains that"   
## [22] "she took the"   
## [23] "leap from"   
## [24] "working in"   
## [25] "hospitality to"   
## [26] "training as an"   
## [27] "apprentice"   
## [28] "because it"   
## [29] "offered far"   
## [30] "better career"   
## [31] "prospects.</p>"   
## [32] "<p>“I didn’t"   
## [33] "know what I"   
## [34] "wanted to do"   
## [35] "when I left"   
## [36] "school, so I"   
## [37] "just started"   
## [38] "working at a"   
## [39] "local cafe,”"   
## [40] "she"   
## [41] "recalls.</p>"   
## [42] "<p>“It was a"   
## [43] "good job, don’t"   
## [44] "get me wrong,"   
## [45] "but I couldn’t"   
## [46] "really see it"   
## [47] "going anywhere."   
## [48] "I wanted to get"   
## [49] "a proper"   
## [50] "qualification"   
## [51] "that would lead"   
## [52] "to a decent"   
## [53] "career.”</p>"   
## [54] "<p>For dentist"   
## [55] "Margaret"   
## [56] "Naylor, Black’s"   
## [57] "experience sums"   
## [58] "up how"   
## [59] "apprenticeships"   
## [60] "empower people"   
## [61] "of any age to"   
## [62] "train for a new"   
## [63] "career. She"   
## [64] "believes they"   
## [65] "open up a wider"   
## [66] "supply of"   
## [67] "cost-effective"   
## [68] "talent to"   
## [69] "employers. It"   
## [70] "is a point that"   
## [71] "the <a"   
## [72] "href=\"https://www.theguardian.com/invest-in-an-apprentice/2018/oct/26/everything-employers-need-to-know-about-apprenticeships\">National"  
## [73] "Apprenticeship"   
## [74] "Service</a>,"   
## [75] "which supports"   
## [76] "England’s"   
## [77] "apprenticeship"   
## [78] "delivery, is"   
## [79] "keen to make"   
## [80] "clear to"   
## [81] "employers:"   
## [82] "apprenticeships"   
## [83] "are not limited"   
## [84] "to school"   
## [85] "leavers.</p>"   
## [86] "<figure"   
## [87] "class=\"element"   
## [88] "element-image"   
## [89] "element--supporting\""   
## [90] "data-media-id=\"53367990a844b511313e6a08113fcfb1f9d1f754\">"   
## [91] "<img"   
## [92] "src=\"https://media.guim.co.uk/53367990a844b511313e6a08113fcfb1f9d1f754/0\_0\_5600\_3360/1000.jpg\""   
## [93] "alt=\"Dentist"   
## [94] "putting on"   
## [95] "surgical"   
## [96] "gloves,"   
## [97] "preparing to"   
## [98] "use dentistry"   
## [99] "equipmentGettyImages-951629888\""   
## [100] "width=\"1000\""   
## [101] "height=\"600\""   
## [102] "class=\"gu-image\""   
## [103] "/> <figcaption>"   
## [104] "<span"   
## [105] "class=\"element-image\_\_caption\">Apprentices"   
## [106] "can find"   
## [107] "themselves"   
## [108] "professionally"   
## [109] "qualified in"   
## [110] "the time it"   
## [111] "takes to"   
## [112] "complete an"   
## [113] "undergraduate"   
## [114] "course.</span>"   
## [115] "<span"   
## [116] "class=\"element-image\_\_credit\">Photograph:"   
## [117] "Eugenio"   
## [118] "Marongiu/Getty"   
## [119] "Images</span>"   
## [120] "</figcaption>"   
## [121] "</figure> <p>In"   
## [122] "fact, there is"   
## [123] "a special"   
## [124] "provision to"   
## [125] "keep the cost"   
## [126] "of recruitment"   
## [127] "low, so an SME,"   
## [128] "such as"   
## [129] "Dental22, which"   
## [130] "is looking to"   
## [131] "balance its"   
## [132] "wage bill, is"   
## [133] "not required to"   
## [134] "match the"   
## [135] "salary an older"   
## [136] "person may have"   
## [137] "been earning in"   
## [138] "another"   
## [139] "position."   
## [140] "Instead, a"   
## [141] "first-year"   
## [142] "apprentice is"   
## [143] "paid the"   
## [144] "minimum wage of"   
## [145] "£3.70 per hour,"   
## [146] "regardless of"   
## [147] "age, minimising"   
## [148] "the burden on"   
## [149] "the"   
## [150] "employer.</p>"   
## [151] "<p>For Naylor,"   
## [152] "though, the"   
## [153] "main benefit"   
## [154] "apprentices"   
## [155] "offer is"   
## [156] "flexibility in"   
## [157] "training people"   
## [158] "to her high"   
## [159] "standards,"   
## [160] "rather than"   
## [161] "hoping that"   
## [162] "qualified staff"   
## [163] "have already"   
## [164] "been trained"   
## [165] "well at another"   
## [166] "practice.</p>"   
## [167] "<p>“Apprenticeships"   
## [168] "are all about"   
## [169] "flexibility."   
## [170] "They allow"   
## [171] "people to"   
## [172] "change jobs and"   
## [173] "qualify in"   
## [174] "another area,"   
## [175] "just as much as"   
## [176] "they help"   
## [177] "school leavers"   
## [178] "get"   
## [179] "professional"   
## [180] "qualifications"   
## [181] "while they"   
## [182] "earn,” she"   
## [183] "says.</p>"   
## [184] "<p>“For us, as"   
## [185] "a business, the"   
## [186] "beauty of"   
## [187] "apprenticeships"   
## [188] "is that we can"   
## [189] "mould employees"   
## [190] "to meet our"   
## [191] "high standards."   
## [192] "Sometimes we"   
## [193] "find that if we"   
## [194] "hire dental"   
## [195] "nurses who have"   
## [196] "qualified"   
## [197] "elsewhere,"   
## [198] "they’re not"   
## [199] "quite up to our"   
## [200] "standards.</p>"   
## [201] "<p>“It means we"   
## [202] "can end up"   
## [203] "having to"   
## [204] "retrain people,"   
## [205] "so for us, it"   
## [206] "just makes a"   
## [207] "lot more sense"   
## [208] "to train our"   
## [209] "own"   
## [210] "apprentices.”</p>"   
## [211] "<aside"   
## [212] "class=\"element"   
## [213] "element-rich-link"   
## [214] "element--thumbnail\">"   
## [215] "<p>"   
## [216] "<span>Related:"   
## [217] "</span><a"   
## [218] "href=\"https://www.theguardian.com/invest-in-an-apprentice/2018/oct/26/many-of-the-partners-at-our-firm-started-as-apprentices\">‘Many"   
## [219] "of the partners"   
## [220] "at our firm"   
## [221] "started as"   
## [222] "apprentices’</a>"   
## [223] "</p> </aside>"   
## [224] "<p>Andi Martin"   
## [225] "is an example"   
## [226] "of a newly"   
## [227] "qualified"   
## [228] "dental nurse"   
## [229] "who has been"   
## [230] "trained to a"   
## [231] "high level"   
## [232] "within Dental22"   
## [233] "over the past"   
## [234] "four years. She"   
## [235] "finished a"   
## [236] "level 3 NVQ in"   
## [237] "dental nursing"   
## [238] "in January,"   
## [239] "having joined"   
## [240] "straight after"   
## [241] "finishing her"   
## [242] "A-levels.</p>"   
## [243] "<p>It was the"   
## [244] "cost of"   
## [245] "university, and"   
## [246] "the size of the"   
## [247] "loan she would"   
## [248] "need to pay,"   
## [249] "that convinced"   
## [250] "Martin to"   
## [251] "become an"   
## [252] "apprentice.</p>"   
## [253] "<p>“I looked in"   
## [254] "to the cost of"   
## [255] "university and"   
## [256] "it would have"   
## [257] "been in the"   
## [258] "£30,000 zone,”"   
## [259] "she says.</p>"   
## [260] "<p>“That’s a"   
## [261] "scary amount of"   
## [262] "money when you"   
## [263] "don’t know for"   
## [264] "sure whether"   
## [265] "there will be a"   
## [266] "job at the end"   
## [267] "of it. So I"   
## [268] "went for the"   
## [269] "option of"   
## [270] "earning while I"   
## [271] "was learning –"   
## [272] "where the"   
## [273] "training is"   
## [274] "paid for you"   
## [275] "and you’re not"   
## [276] "left with a"   
## [277] "pile of"   
## [278] "debt.”</p>"   
## [279] "<p>Taking the"   
## [280] "apprenticeship"   
## [281] "route meant she"   
## [282] "would be"   
## [283] "professionally"   
## [284] "qualified by"   
## [285] "the time she"   
## [286] "would have been"   
## [287] "graduating."   
## [288] "She’d also be"   
## [289] "without debt"   
## [290] "and already"   
## [291] "employed.</p>"   
## [292] "<p>In Naylor’s"   
## [293] "experience,"   
## [294] "training up an"   
## [295] "apprentice,"   
## [296] "such as Martin,"   
## [297] "not only"   
## [298] "benefits her"   
## [299] "business but"   
## [300] "also the wider"   
## [301] "dentistry and"   
## [302] "healthcare"   
## [303] "industry,"   
## [304] "because many"   
## [305] "apprentices"   
## [306] "carry on their"   
## [307] "studies,"   
## [308] "meaning the"   
## [309] "sector has a"   
## [310] "constant flow"   
## [311] "of"   
## [312] "higher-qualified"   
## [313] "staff.</p>"   
## [314] "<p>“A lot of"   
## [315] "apprentices"   
## [316] "qualify as"   
## [317] "dental nurses"   
## [318] "and then go on"   
## [319] "to study to"   
## [320] "become dental"   
## [321] "therapists,”"   
## [322] "she says.</p>"   
## [323] "<p>“A therapist"   
## [324] "can do a lot"   
## [325] "more work than"   
## [326] "a nurse: they"   
## [327] "can do"   
## [328] "restorations,"   
## [329] "fillings, take"   
## [330] "impressions and"   
## [331] "remove"   
## [332] "children’s"   
## [333] "teeth."   
## [334] "Sometimes,"   
## [335] "dental nurses"   
## [336] "go on to become"   
## [337] "therapists or"   
## [338] "take other"   
## [339] "roles within"   
## [340] "the NHS.</p>"   
## [341] "<p>“By becoming"   
## [342] "apprentices and"   
## [343] "getting"   
## [344] "qualified,"   
## [345] "young people"   
## [346] "get access to"   
## [347] "these new"   
## [348] "career paths,"   
## [349] "and the"   
## [350] "industry gets a"   
## [351] "steady flow of"   
## [352] "highly"   
## [353] "qualified"   
## [354] "staff.”</p>"

print("Sample of Clean News Article Body:")

## [1] "Sample of Clean News Article Body:"

cat("\n")

news$body[2]

## [1] "sitting in a plush hotel lobby in sunny southern spain matt white is unusually calm the head sports director of world tour cycling team mitchelton scott is ordinarily a frenetic ball of energy on race day finalising logistics plotting team tactics and counselling his riders on their duties for the stage ahead but during the two thousand eighteen vuelta a espana which began on saturday white appears downright relaxed indeed when he sits down for an interview with guardian australia it is rugby not cycling that is at the forefront of his mind having just watched australias latest loss to the all blacks this atypical serenity is partly because after guiding mitchelton scott through the gruelling giro ditalia and tour de france the forty four year old australian is taking a well deserved break he is in spain to mentor another sports director rather than being in the hot seat himself but it is also because unlike many other teams at the vuelta mitchelton scotts season does not depend on a strong finale we are three wins away from this being our most successful season he says having finished his dissection of the wallabies and we still have two months left related yates twins reunited for simons vuelta a espana glory quest there is every chance that the weeks ahead will bring more success after a break out giro where he won three stages and wore the pink jersey for thirteen stages britains simon yates is among the favourites for the tour of spain with team skys chris froome and geraint thomas absent richie portes fitness uncertain and vincenzo nibali reportedly focusing on the world championships yates has a golden opportunity to seize his and mitchelton scotts first grand tour it is an open vuelta says white we want to win it twenty six year old yates is in good form after a recent tour de pologne stage win and came to spain backed by a strong squad including his twin brother adam simon has really stamped his name this year as one of the best riders in the world he is just one step away from arriving on the podium adam yates on stage two of the vuelta between marbella and caminito del rey photograph michael steele getty images whites two thousand eighteen has not been entirely smooth sailing though mitchelton scott drew heavy criticism when they omitted prodigious young australian sprinter caleb ewan from their tour de france roster despite having foreshadowed his debut at the race six months prior ewan complained on twitter that devastated is an understatement and the non selection seemingly hastened his decision to move elsewhere the twenty four year old subsequently signed with lottosoudal for two thousand nineteen i havent spoken to caleb since i rang him and told him he was not doing the tour de france says white when asked if bad blood lingered our race calendars have not coincided people try to make a big deal about it but at the end of the day this is professional sport this is business people make decisions about changing the roster about tactics you do not see people complaining in football codes when they get left on the bench ewan was omitted to enable mitchelton scott to focus on adam yates yellow jersey campaign in france which ultimately faltered it is hard to support for both general classification and sprints white explains but his preferencing of an englishman over a sydneysider attracted the ire of some commentators who claimed the australian registered team had lost its much lauded australian dna white is visibly bothered by the suggestion no i do not think that is a fair criticism he replies if you look at the breakdown of the team in two thousand twelve we had thirty riders and seventeen australians now we have twenty five and eleven the difference now is that some of our leaders are not australian the australian dna has always been the same the sports director cites fc barcelona they have catalans in the team and they represent catalonia but leo messi is argentinian does it matter you are following us because of who we are and what we stand for not because of how many australians we have in the line up it is frustrating when you see how we are judged these ultimately are minor gripes it is just seven years since mitchelton scott began from scratch and in that time they have finished on the podium at two grand tours claimed four of the five one day monuments and won a handful of iconic week long races having recently resigned the yates brothers and with esteban chaves still under contract white has a plethora of general classification options at his disposal while young australian talents jack haig robert power and lucas hamilton wait in the wings the mitchelton scott womens team meanwhile goes from strength to strength and this year won the iconic giro rosa we are one of the best teams in the world white says we are competitive everywhere we go we are knocking on the door of a grand tour victory and i am confident that we can win one soon if i walked away tomorrow i would be very proud but there is unfinished business"