

KevinFijiAgreementOnNegationAnnotation.Rmd

Inter-annotator agreement between Kevin and Fiji on CRAFT

and MIMIC2MD negation annotation

```
craft.kevin.data <- read.table("/Users/kev/Dropbox/Negation/Code/CRAFT.type.token.with.prefix.KEVIN.csv")
craft.fiji.data <- read.table("/Users/kev/Dropbox/Negation/Code/CRAFT.type.token.with.prefix.FIJI.csv",
mimic.kevin.data <- read.table("/Users/kev/Dropbox/Negation/Code/MIMIC2MD.type.token.with.prefix.KEVIN.csv")
mimic.fiji.data <- read.table("/Users/kev/Dropbox/Negation/Code/MIMIC2MD.type.token.with.prefix.FIJI.csv")
```

Let's define a function to calculate agreement. To test it, I calculated my agreement with myself and got 1.0, then my agreement on the two different corpora and got a surprisingly high 0.06!

```
agreement <- function(df.one, df.two) {
  # quick validation
  # this doesn't quit when it should--TODO how do I do this in R??
  if (nrow(df.one) != nrow(df.two)) { quit }

  agreements <- 0
  disagreements <- 0
  #judgements <- 0 # calculating agreement only when one or the other
  # labelled it
  for (i in 1:nrow(df.one)) {
    if (df.one[i,4] == "x" & df.two[i,4] == "x") {
      agreements <- agreements + 1
    } else if ((df.one[i,4] == "x" & df.two[i,4] != "x") | (df.one[i,4] != "x" & df.two[i,4] == "x")) {
      disagreements <- disagreements + 1
    }
  }
  return (agreements / (agreements + disagreements))
}
```

Let's define another function to calculate agreement, this one considering the ones that neither of us annotated as negative. To test it, I calculated my agreement with myself and got 1.0.

```
agreement.including.all <- function(df.one, df.two) {
  # quick validation
  # this doesn't quit when it should--TODO how do I do this in R??
  if (nrow(df.one) != nrow(df.two)) { quit }

  agreements <- 0
  disagreements <- 0
  #judgements <- 0 # calculating agreement only when one or the other
  # labelled it
  for (i in 1:nrow(df.one)) {
    if (df.one[i,4] == df.two[i,4]) {
      agreements <- agreements + 1
    } else {
      disagreements <- disagreements + 1
    }
  }
}
```

```

    }
  }
  return (agreements / (agreements + disagreements))
}

```

Now let's try to calculate agreement!

```

agreement.craft <- agreement(mimic.kevin.data, mimic.fiji.data)
print("CRAFT IAA (only including ones marked by someone):")

```

```
## [1] "CRAFT IAA (only including ones marked by someone):"
```

```
agreement.craft
```

```
## [1] 0.6785714
```

```

agreement.mimic <- agreement(craft.kevin.data, craft.fiji.data)
print("MIMIC IAA (as above):")

```

```
## [1] "MIMIC IAA (as above):"
```

```
agreement.mimic
```

```
## [1] 0.766443
```

```

agreement.craft <- agreement.including.all(mimic.kevin.data, mimic.fiji.data)
print("CRAFT IAA (including all):")

```

```
## [1] "CRAFT IAA (including all):"
```

```
agreement.craft
```

```
## [1] 0.9338583
```

```

agreement.mimic <- agreement.including.all(craft.kevin.data, craft.fiji.data)
print("MIMIC IAA (as above):")

```

```
## [1] "MIMIC IAA (as above):"
```

```
agreement.mimic
```

```
## [1] 0.9471285
```

```
all.kevin.data <- rbind(craft.kevin.data, mimic.kevin.data)
all.fiji.data <- rbind(craft.fiji.data, mimic.fiji.data)
agreement.all <- agreement.including.all(all.kevin.data, all.fiji.data)
print("Overall agreement, including all:")
```

```
## [1] "Overall agreement, including all:"
```

```
agreement.all
```

```
## [1] 0.9422633
```