**Question 1 - multiple choice, no shuffle**

Gender biases were previously shown to exist in the Word2Vec embedding by which paper?

A: Angwin et al. (2016)

\*B: Bolukbasi et al. (2016)

C: Caliskan et al. (2017)

D: Dehghan et al. (2017)

**Question 2 - numeric**

How many unique individuals are presented in the phenotypically balanced dataset that is provided by this paper?

\*A: 1270.0

Default Feedback:

**Question 3 - text match**

What is the name of the phenotypically balanced data set that is provided by this paper?

\*A: The Pilot Parliaments Benchmark

\*B: Pilot Parliaments Benchmark

\*C: PPB

Default Feedback:

**Question 4 - multiple choice, shuffle**

Gender classification accuracy is evaluted using intersectional subgroups of:

A: Gender and Ethnicity

B: Gender and Nationality

C: Gender and Race

\*D: Gender and Skin Type

**Question 5 - checkbox, shuffle, partial credit**

Which of the following are commercial gender classification systems used in the analysis within this paper?

A: Adience

B: Google

\*C: F​ace++

D: A​ffectiva

E: F​aception

\*F: M​iscrosoft

\*G: I​BM

**Question 6 - multiple choice, no shuffle**

The gender classification systems used in this paper use training data labelled with:

A: G​ender Identity

B: B​iological Sex

C: B​oth

D: N​either

\*E: U​nknown

**Question 7 - multiple choice, shuffle**

The positive predictive value (PPV) is used as an accuracy metric in this paper. Which of the following is definitions relates to PPV?

\*A: number of individuals correctly labelled / total number of individuals

B: number of individuals correctly labelled / number of individuals incorrectly labelled

C: total number of individuals / number of individuals correctly labelled

D: number of individuals incorrectly labelled / number of individuals correctly labelled

**Question 8 - multiple choice, no shuffle**

Which sub-group has the highest true positive rate across all classification systems?

A: D​arker Females

B: Darker Males

\*C: L​ighter Females

D: L​ighter Males

**Question 9 - multiple choice, no shuffle**

Which sub-group has the highest error rate across all classification systems?

\*A: D​arker Females

B: D​arker Males

C: L​ighter Females

D: L​ighter Males

**Question 10 - multiple choice, no shuffle**

Which sub-group has the highest postive predictive value for the Face++ classification system?

A: D​arker Females

\*B: D​arker Males

C: L​ighter Females

D: L​ighter Males