

Items from Psychometric Tests as Training Data for Personality Profiling Models of Twitter Users

Anne Kreuter, Kai Sassenberg, and Roman Klinger

(1) Motivation

Data for psychological concept modelling in social media can be obtained by:

1. Users fill questionnaire and share their profile
2. Rely on self reports of users
3. Annotators annotate the concept of users

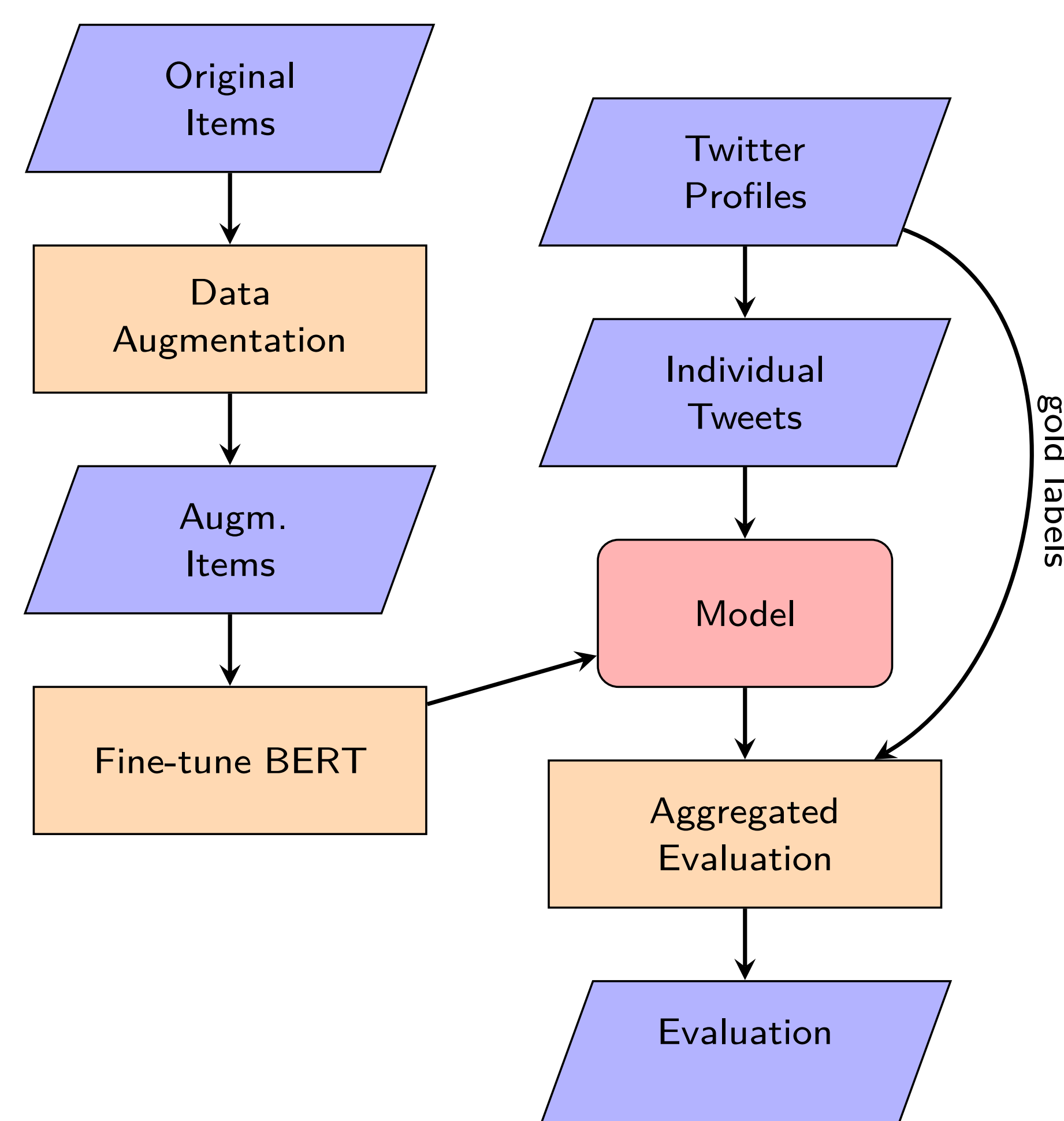
These approaches have issues regarding data privacy and reliability. Solution:

⇒ Train models directly on questionnaires.

(2) IPIP Neo 300 Test Examples

Variable	Cor.	Item Text
Openness	+	“Am interested in many things.”
Openness	-	“Do not like art.”
Conscientiousness	+	“Know how to get things done.”
Conscientiousness	-	“Find it difficult to get down to work.”
Extraversion	+	“Warm up quickly to others.”
Extraversion	-	“Am hard to get to know.”
Agreeableness	+	“Am concerned about others”
Agreeableness	-	“Can’t stand weak people”
Neuroticism	+	“Am afraid of many things.”
Neuroticism	-	“Feel comfortable with myself.”

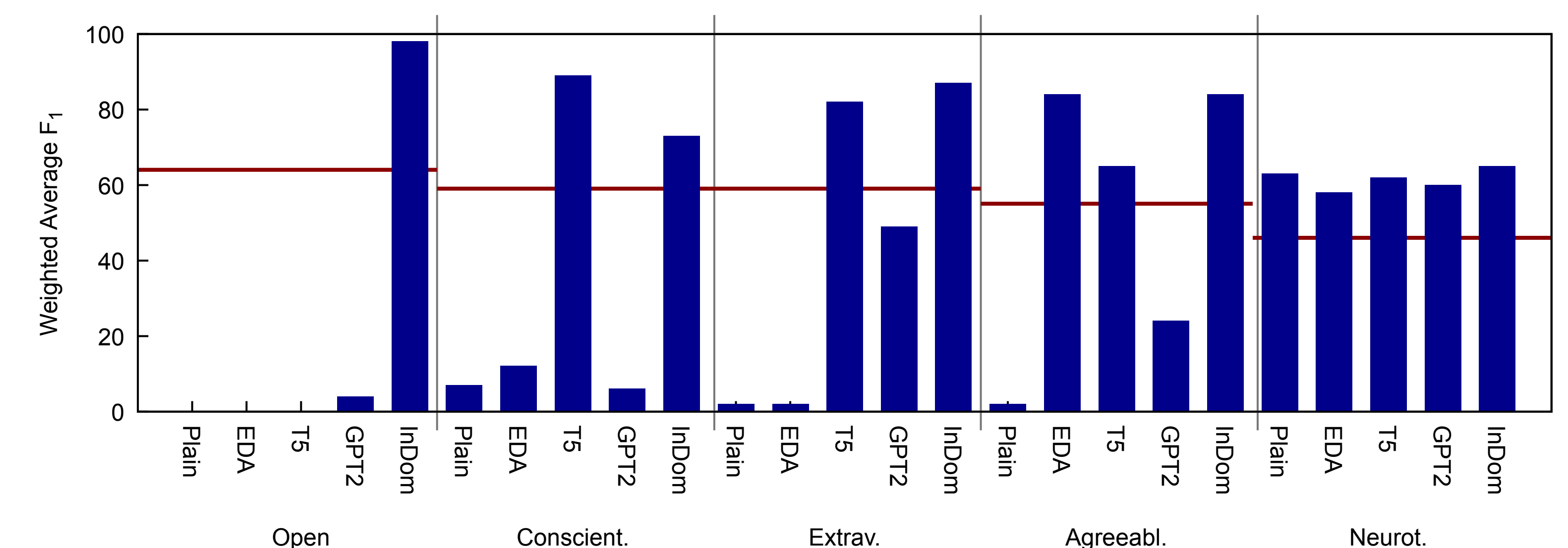
(3) Workflow



(4) Corpora

Class.	PAN 2015					
	IPIP-NEO		Profiles		Tweets	
	+	-	+	-	+	-
Open.	28	32	288	3	26,743	236
Consc.	31	29	229	15	21,391	1,428
Extra.	36	24	235	21	21,686	2,000
Agree.	24	36	223	29	20,441	2,831
Neurot.	33	27	76	197	18,076	7,168

(5) Results



(6) Introspection

T Tweet

O @username What my love life will hold instore for me in the future. I'd never ask when I'm gonna die...?????????

E @username Slade!!! Cool memories of my grammar school days!!

A @username I rocked so much to their music!

(7) Conclusion

- Model works comparably well for a subset of traits
- Introspection: decisions based on reasonable concepts
- Promising route to get less biased classifiers
- Future Work: Test this approach across many concepts