Saint or Sinner? Language-action Cues for Modeling Deception Using Support Vector Machines

SHUYUAN MARY HO
XIUWEN LIU
CHERYL BOOTH
ARAVIND HARIHARAN
FLORIDA STATE UNIVERSITY – SCHOOL OF INFORMATION

INTRODUCTION

CMC-based technologies

- ✓Increase the speed, geographical scope, and convenience of interpersonal communication.
- ✓Increase users' exposure to deceptive online communication (e.g., phishing) and its attendant risks (e.g., fraud, identity theft).

Challenges/Success factors:

- ✓Evaluate the truthfulness of a statement
- ✓Evaluate the trustworthiness of a communicating party
- ✓Evaluate the authenticity and creditability of the information exchanged

"Cue lean" text-based CMC makes intent particularly difficult to discern

OBJECTIVE and IMPACT

- Finding the optimal modeling approach, and identifying those language-action cues most indicative of deceptive intent, to inform the development of a machine learning system to automatically detect deception in text-based CMC.
- Can be applied to analysis of transcribed text
  - e.g. emergency (911)-call transcripts; other VOIP and similar communications.
  - Eventual inclusion of voice cues?
- An “online polygraph”.

RESEARCH QUESTION

Can we computationally classify deception in spontaneous computer-mediated communication across a pluralistic background of users?

"Language-action cues" refers to linguistic styles, phrases, patterns, or actions in an actor's written expression and manifested as an indirect or subtle signal to other actors (Ho, Hancock et al. 2016, 2015).

DECEPTION

Two fundamental truths

1. Humans are not good at detecting deception (Ekman 1991).
2. Deception is fairly common, occuring in approx. one-quarter of all communications (Buller & Burgoon 1996).

Deception is a volitional and intentional act, that can be revealed through these cues

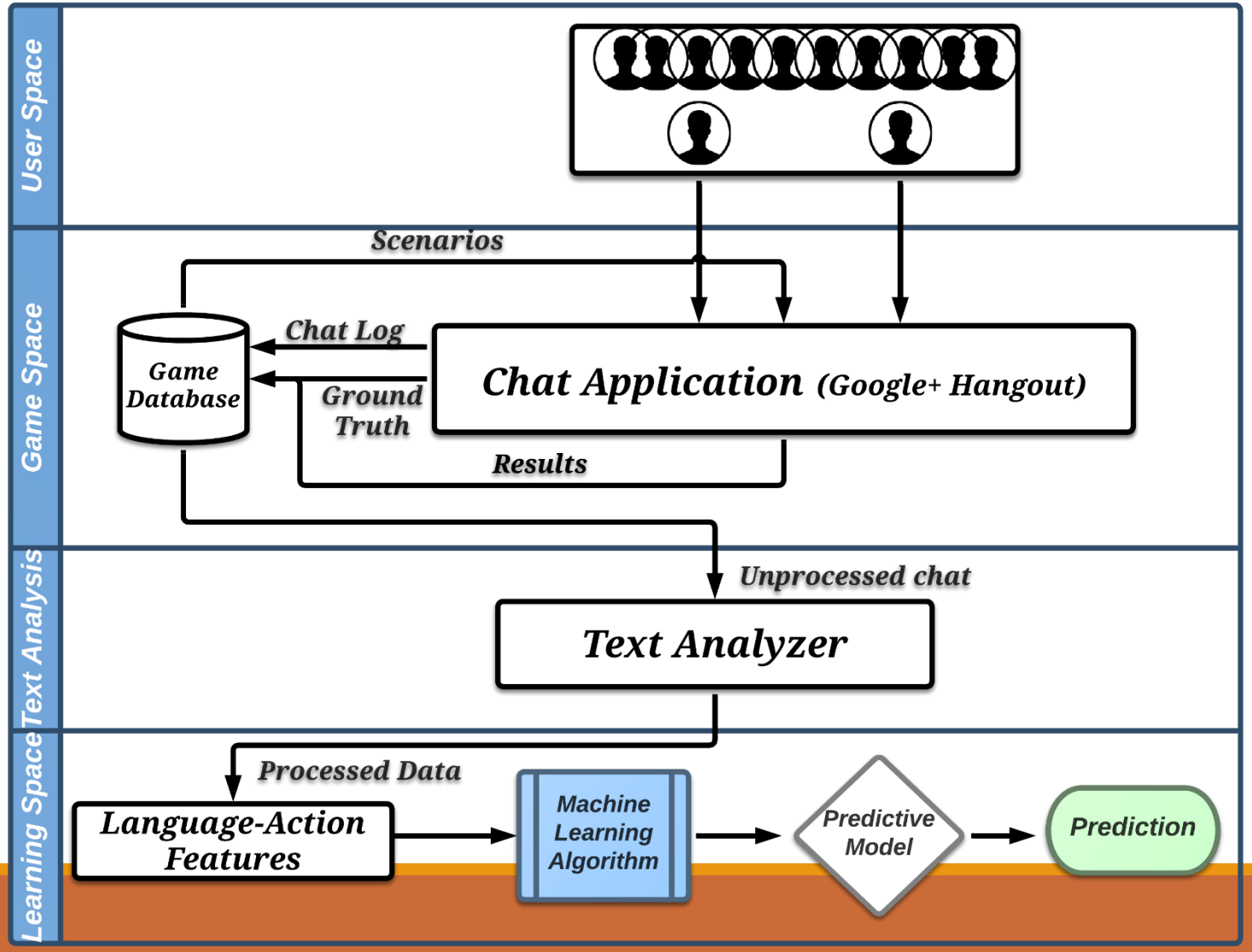
1. IT IS PERVASIVE, STRATEGIC AND CONTEXT-SENSITIVE (Buller & Burgoon, 1996; Ekman et al., 1991, 1969)
2. Planned or “on the fly” (Whitty, 2012)
3. Consequences: little white lie vs. significant (Hancock et al., 2009)
4. Mode of communication: synchronous (Zhou et al., 2008) vs. asynchronous (Zhou et al., 2003)
5. To avoid social discomfort (DePaulo, Kashy et al., 1996; Hoffman et al., 1996)
6. To choose rich media by sending conflicting cues (Daft et al., 1987; Trevino et a., 1987)

Language-action Cues in CMC

- Research has shown that deceptive intent can be revealed through subtle language-action cues/ features.
  - Linguistic Inquiry and Word Count tool
- We frequently miss these cues.
  - We are very poor human sensors individually.
- An automated process or system for identifying these cues would emphasize potential cues to the receiver.
- Leveraging machine learning algorithms can help in developing such a system

LIWC Categories		CODING SCHEMA	Examples
Affective Process		affect	happy, cried, abandon
	Positive Emotion	posemo	love, nice, sweet
	Negative Emotion	negemo	hurt, ugly, nasty
Cognitive Process		cogmech	cause, know, ought
	Certainty	certain	always, never
	Inclusive	incl	and, with, include
	Exclusive	excl	but, without, exclude
	Discrepancy	discrep	should, would, could
	Insight	insight	think, know, consider
	Causation	cause	because, effect, since
	Tentative	tentat	maybe, perhaps, guess
	Inhibition	inhib	block, constrain, stop
Negations		negate	no, not, never
Pronouns		pronoun	
	1st person singular	self-reference	I, me, myself
	1st person plural	self-reference	we, us, our, ours
	2nd person	other reference	you
Word Count		WC	n/a

RESEARCH DESIGN



- ✓Online game simulates real-time interactive scenarios in synchronous communication channels.
- ✓Hosted on Google + Hangout
- ✓Game sessions recorded and stored using MySQL database.

iSensor Lab


The iSensor Research Lab focuses on communication and information behaviors in computer-mediated communication environments that relate to cybersecurity and privacy. Our research objectives are to better understand trusted human-computer interactions as well as factors that promote or interfere with collaboration and communication in virtual organizations. Through a sociotechnical lens, we create and implement social, scenario-based games as experiments that allow us to build and test models and theories of online trusted and deceptive human behavior. We believe that rigorous research can and should be performed in a fun learning environment.
FUN HOUSE
CYBER LAB
Real or Spiel? is an interpersonal synchronous game developed on Google+ Hangout platform. The players' objective is to identify whether their communicating partner is a saint or a sinner.
Collabo is a group-based synchronous game developed on Google+ Hangout platform. The objective is to solve the puzzles and win the game in the overall competition.
Guess Who? is an asynchronous game developed on a standalone platform. Players deposit their statements into a MySQL database. The database randomly displays other players' statements. The main objective is to guess the gender of each statement author.
Whodunit? is a murder mystery game developed on a standalone platform. Players will try to identify a murderer by reviewing a crime scene.
Cybersecurity Forum is a place where we can collect ideas about online games from the members of the Funhouse.
Turing game is an earlier version of the Guess Who? game. Instead of having computers guess the gender of the speaker, players will tag the gender to the statement. The score of the tagging is instantaneous, and displayed in real time.
We thank the following institutions for their support!
Florida State University
CCI
National Science foundation

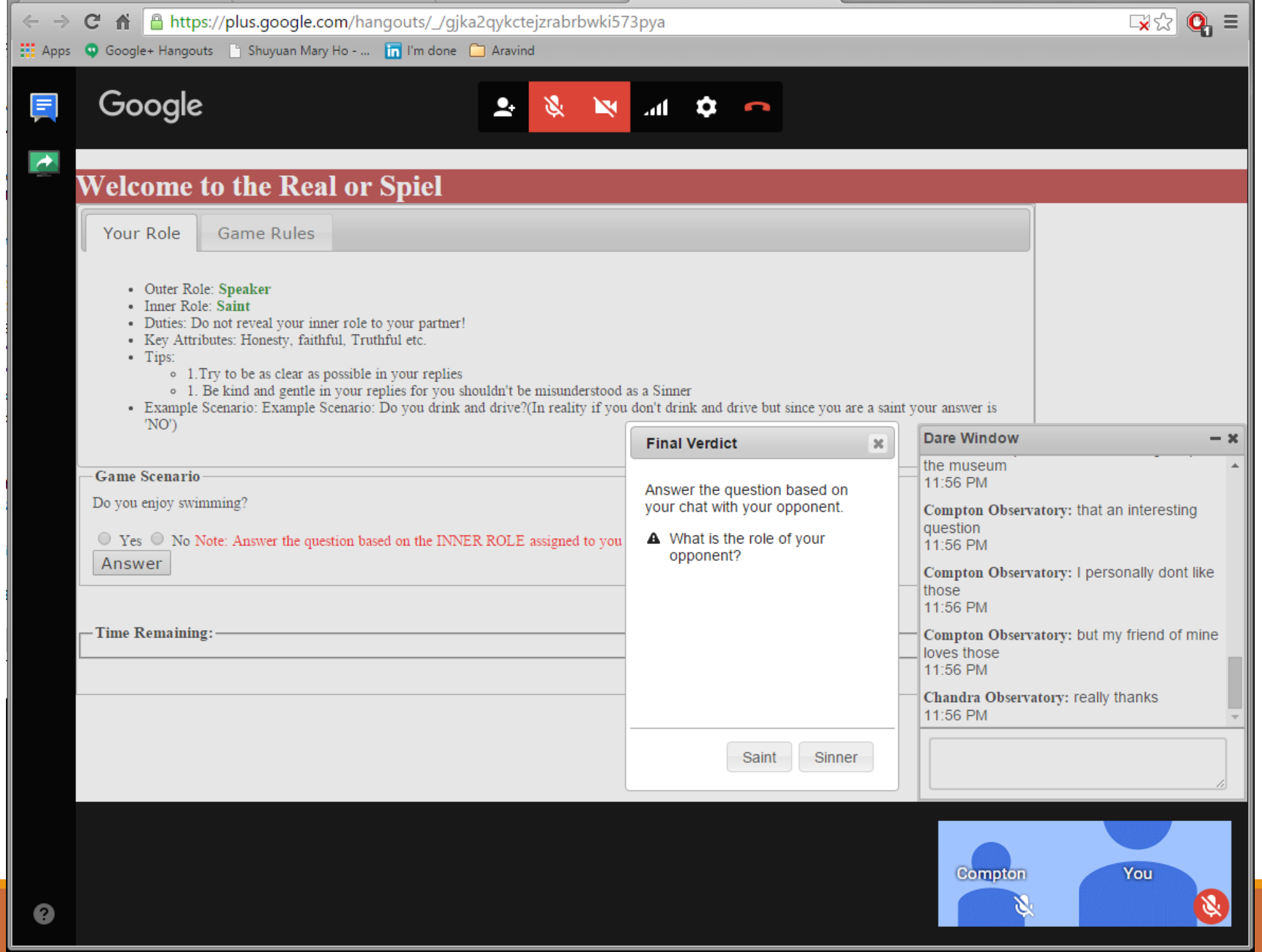
BASIC GAME PLAY

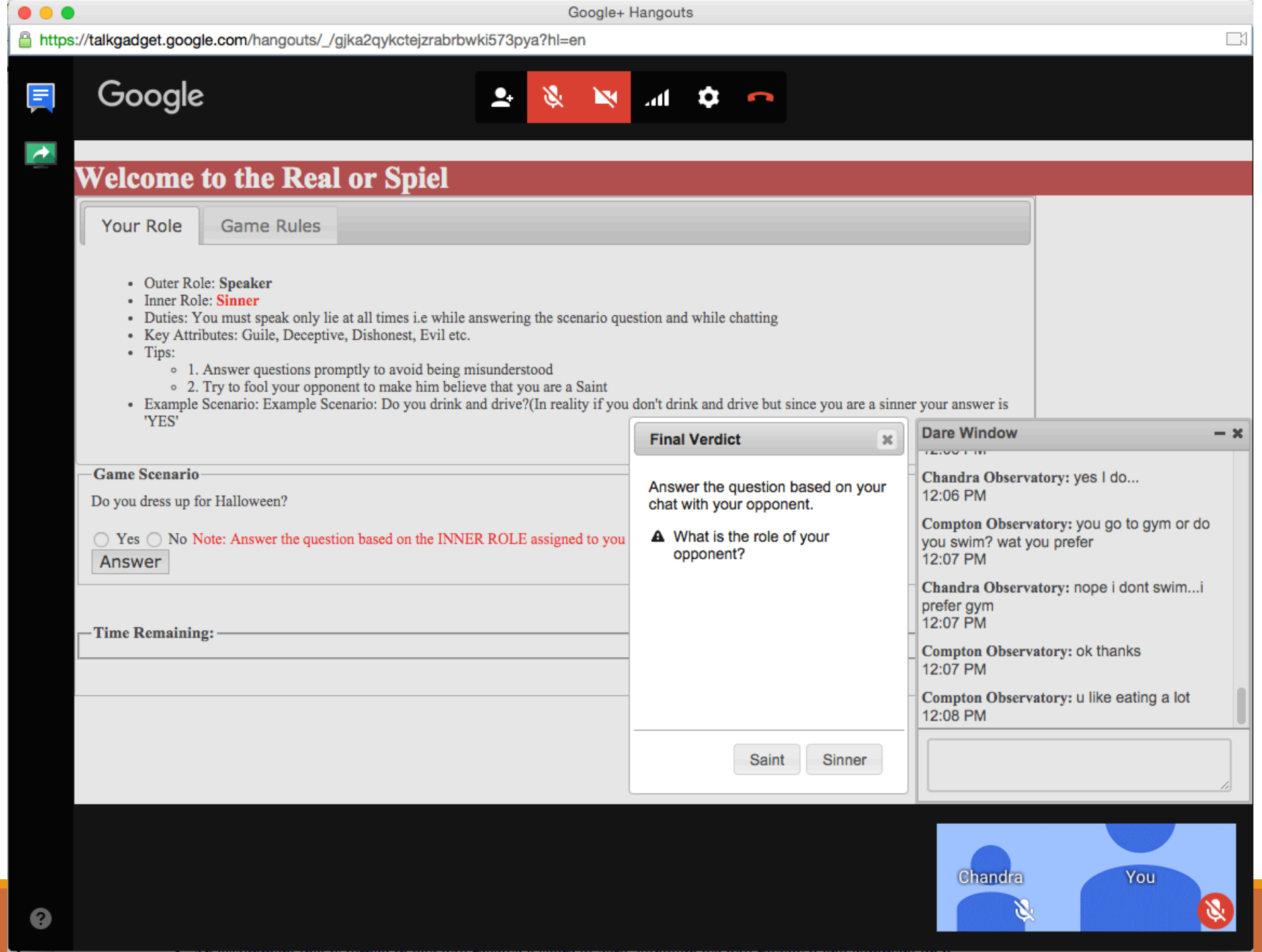
Players are assigned in pairs
- Randomly assigned outer roles as “speaker” or “detector”
- Speakers are also randomly assigned an “inner role”
  - Saint = truthful
  - Sinner = deceptive

Each game session lasts approximately 30 minutes
- Consists of multiple scenarios, with players alternately taking roles as detector or speaker.

Each scenario involves discussion about a specific, random question.
- Starts with the speaker establishing the “ground truth” by answering the question truthfully, “for the record” (i.e. the system) WITHOUT REFERENCE TO INNER ROLE.
  - Response is NOT visible to the detector.
  - Baseline against which the speaker’s responses to the detector’s questions is measured.

Detector asks probing questions to determine the answer to the question.
- Speaker responds to the detector’s questions in accordance with his/her inner role.


https://plus.google.com/hangouts/_/gjka2qykctejzrabrbwki573pya
Apps
Google+ Hangouts
Shuyuan Mary Ho - ...
I'm done
Aravind
Google
Welcome to the Real or Spiel
Your Role
Game Rules
Outer Role: Speaker
Inner Role: Saint
Duties: Do not reveal your inner role to your partner!
Key Attributes: Honesty, faithful, Truthful etc.
Tips:
1.Try to be as clear as possible in your replies
1. Be kind and gentle in your replies for you shouldn't be misunderstood as a Sinner
Example Scenario: Example Scenario: Do you drink and drive?(In reality if you don't drink and drive but since you are a saint your answer is 'NO')
Game Scenario
Do you enjoy swimming?
Yes
No
Note: Answer the question based on the INNER ROLE assigned to you
Answer
Time Remaining:
Final Verdict
Answer the question based on your chat with your opponent.
What is the role of your opponent?
Saint
Sinner
Dare Window
the museum
11:56 PM
Compton Observatory: that an interesting question
11:56 PM
Compton Observatory: I personally dont like those
11:56 PM
Compton Observatory: but my friend of mine loves those
11:56 PM
Chandra Observatory: really thanks
11:56 PM
Compton
You


Google+ Hangouts
https://talkgadget.google.com/hangouts/_/gjka2qykctejzrabrbwki573pya?hl=en
Google
Welcome to the Real or Spiel
Your Role
Game Rules
Outer Role: Speaker
Inner Role: Sinner
Duties: You must speak only lie at all times i.e while answering the scenario question and while chatting
Key Attributes: Guile, Deceptive, Dishonest, Evil etc.
Tips:
1. Answer questions promptly to avoid being misunderstood
2. Try to fool your opponent to make him believe that you are a Saint
Example Scenario: Example Scenario: Do you drink and drive?(In reality if you don't drink and drive but since you are a sinner your answer is 'YES'
Game Scenario
Do you dress up for Halloween?
Yes No Note: Answer the question based on the INNER ROLE assigned to you
Answer
Time Remaining:
Final Verdict
Answer the question based on your chat with your opponent.
What is the role of your opponent?
Saint
Sinner
Dare Window
Chandra Observatory: yes I do...
12:06 PM
Compton Observatory: you go to gym or do you swim? wat you prefer
12:07 PM
Chandra Observatory: nope i dont swim...i prefer gym
12:07 PM
Compton Observatory: ok thanks
12:07 PM
Compton Observatory: u like eating a lot
12:08 PM
Chandra
You

DATA COLLECTION

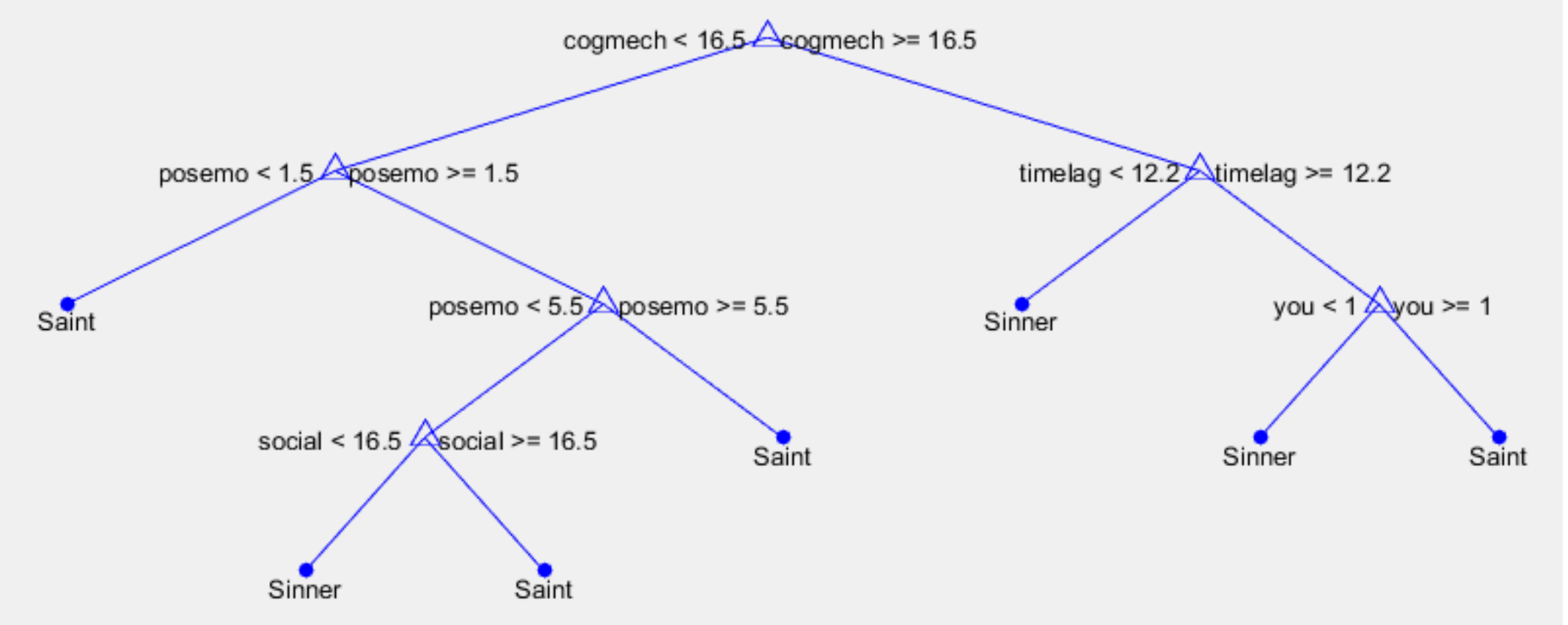
- ✓Data was collected in Fall 2014 and Spring 2015
- ✓80 game sessions with 40 participants; 22 males and 18 females (20 pairs in total) with ages from 18 to 65 years old. Players were replaced with pseudo-names for privacy protection.

Phase/Player	Player 1	Player 2
Phase 1	Speaker & Saint	Detector
Phase 2	Detector	Speaker & Saint
Phase 3	Speaker & Sinner	Detector
Phase 4	Detector	Speaker & Sinner

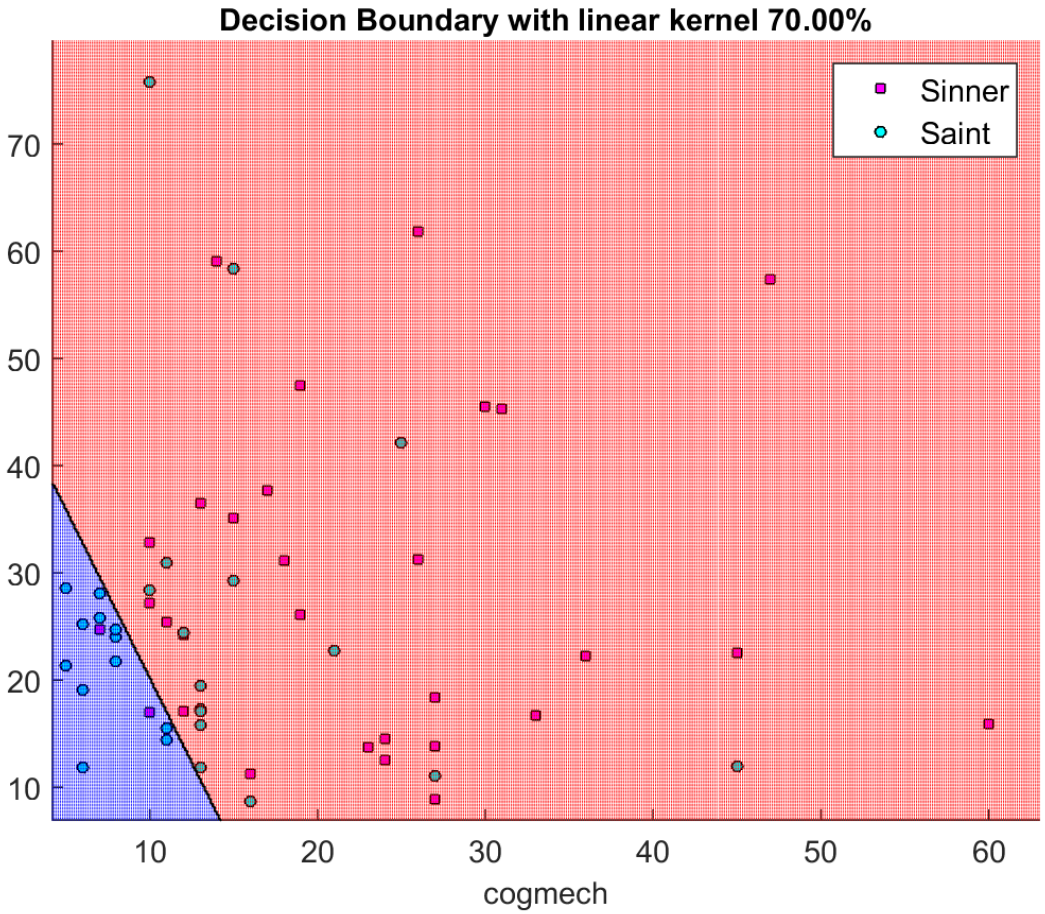
DATA ANALYSIS

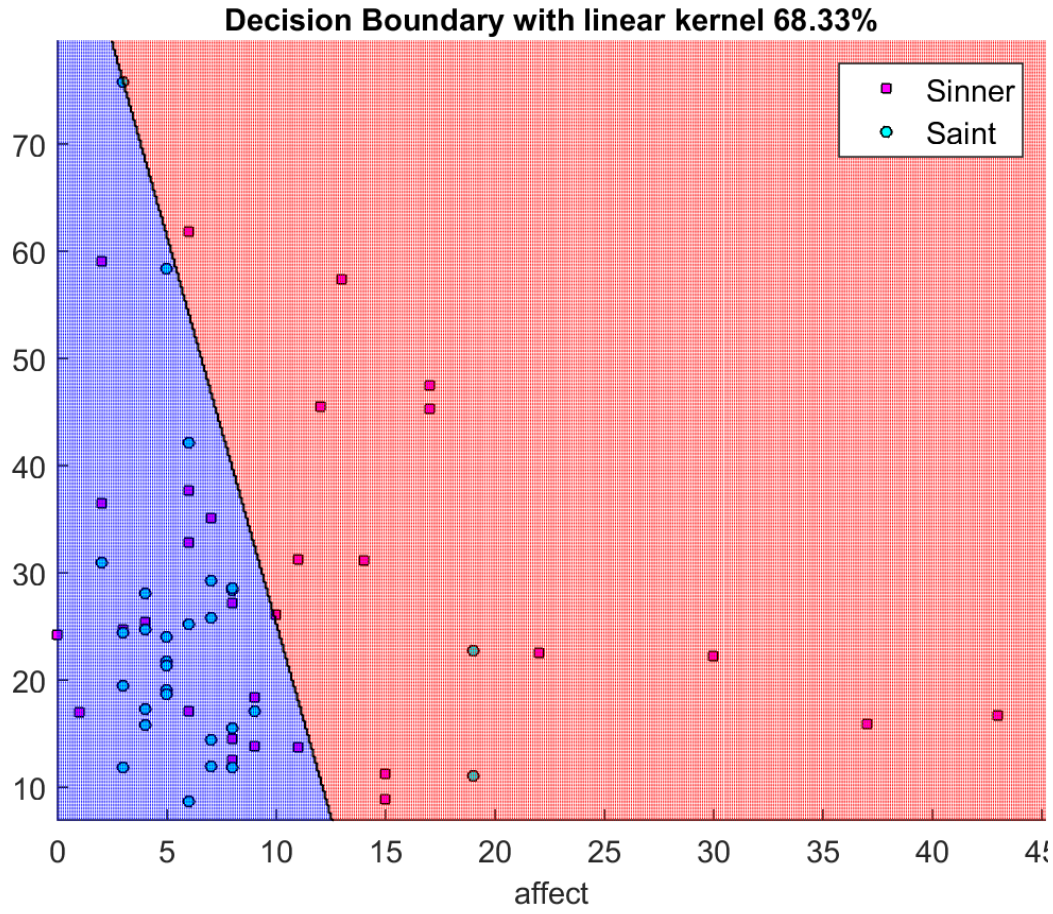
- Across eighty (80) game sessions, forty (40) participants made 1,210 deception guesses/ attempts, yielding a success rate of **52.4%**. This result confirms Ekman and O'Sullivan's (1991) propositions that humans are poor lie detectors, and generally have an accuracy rate of spotting deception at around 50% (nearly random) chances.
- Our final dataset consists of 2,196 lines of chat with a total of 7,271 words, further analyzed by Matlab R2015a.

DECISION TREE Learning

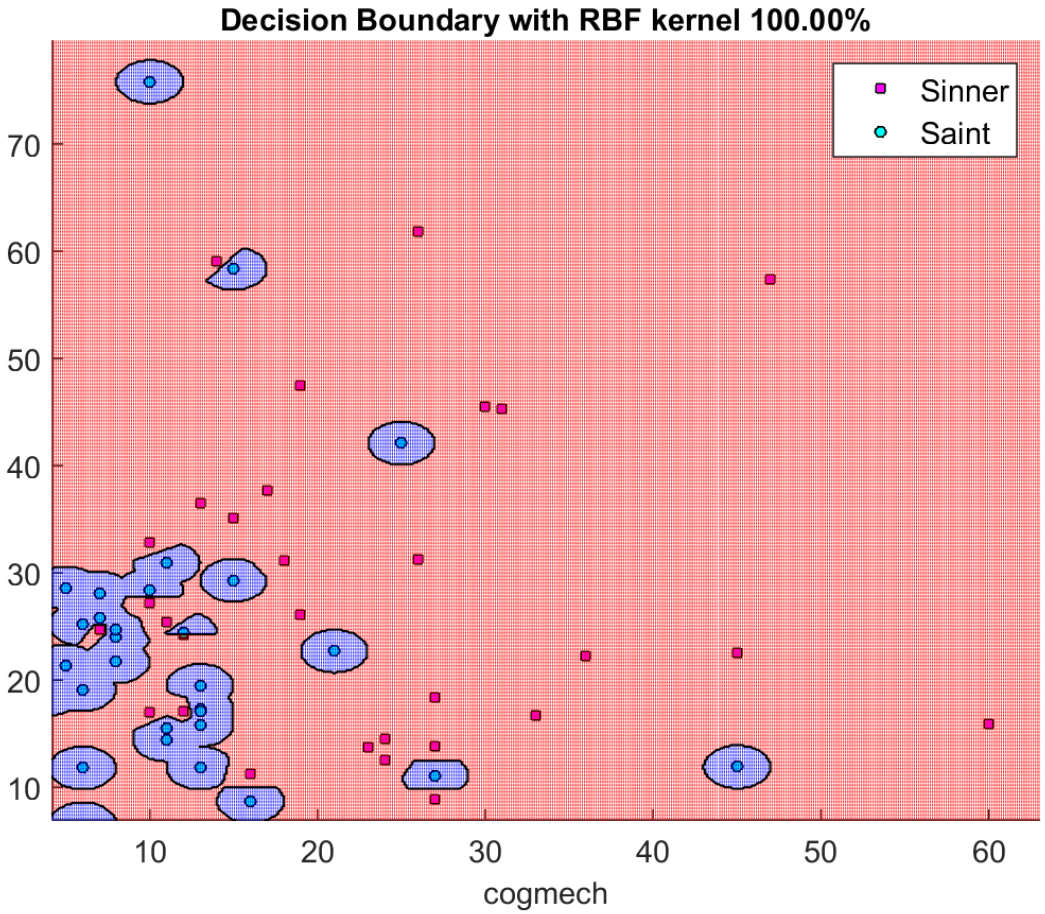


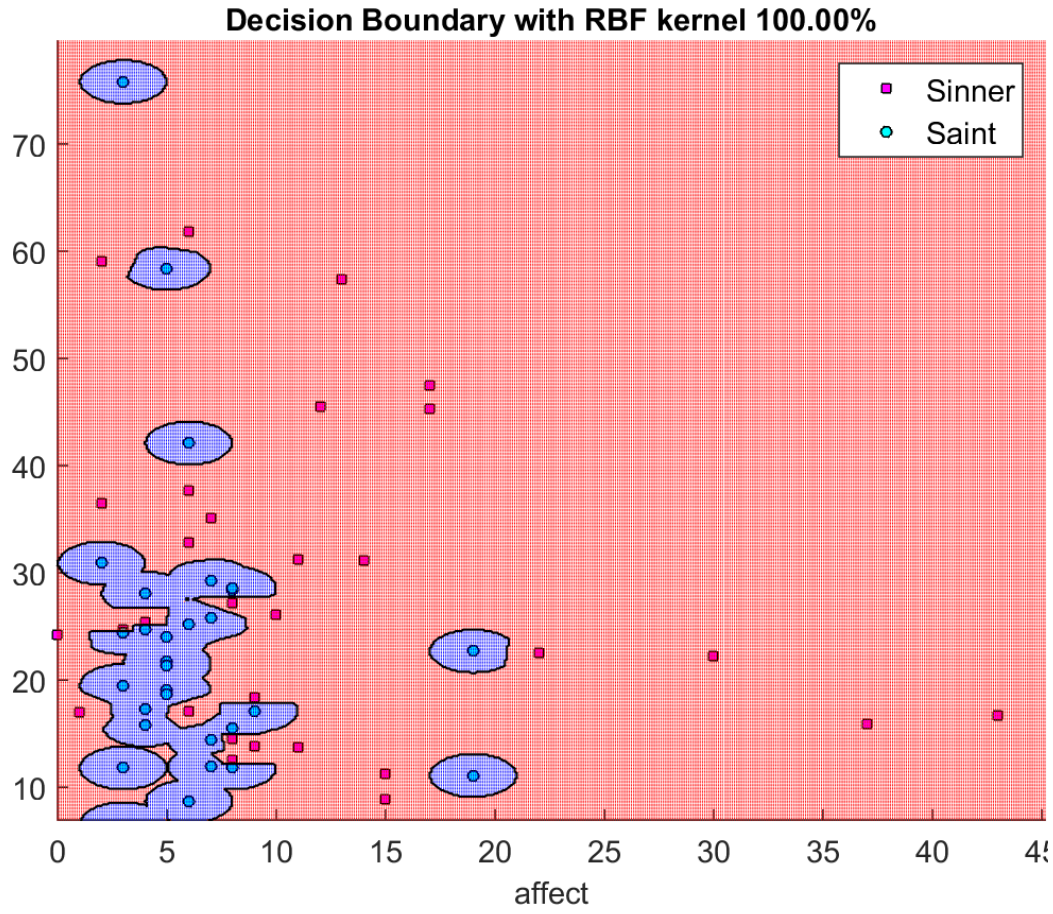
2D Support Vector Machine (SVM) Linear Kernel



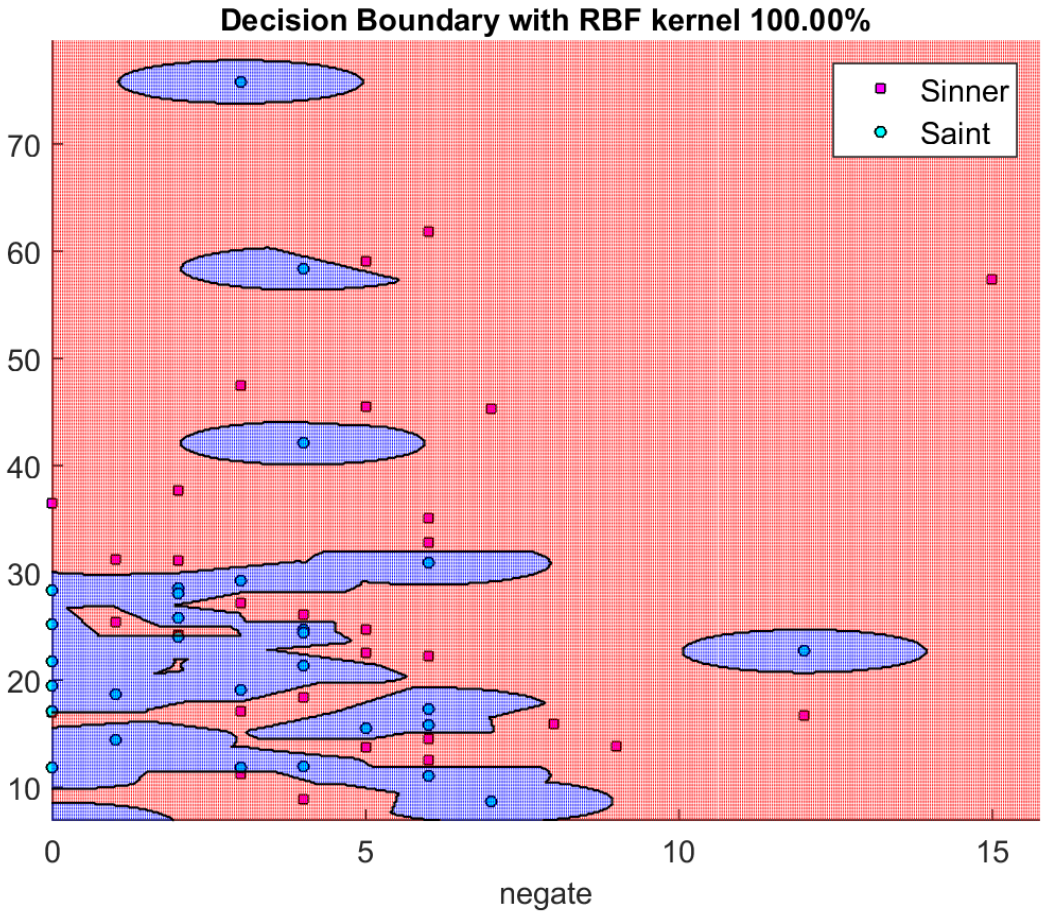


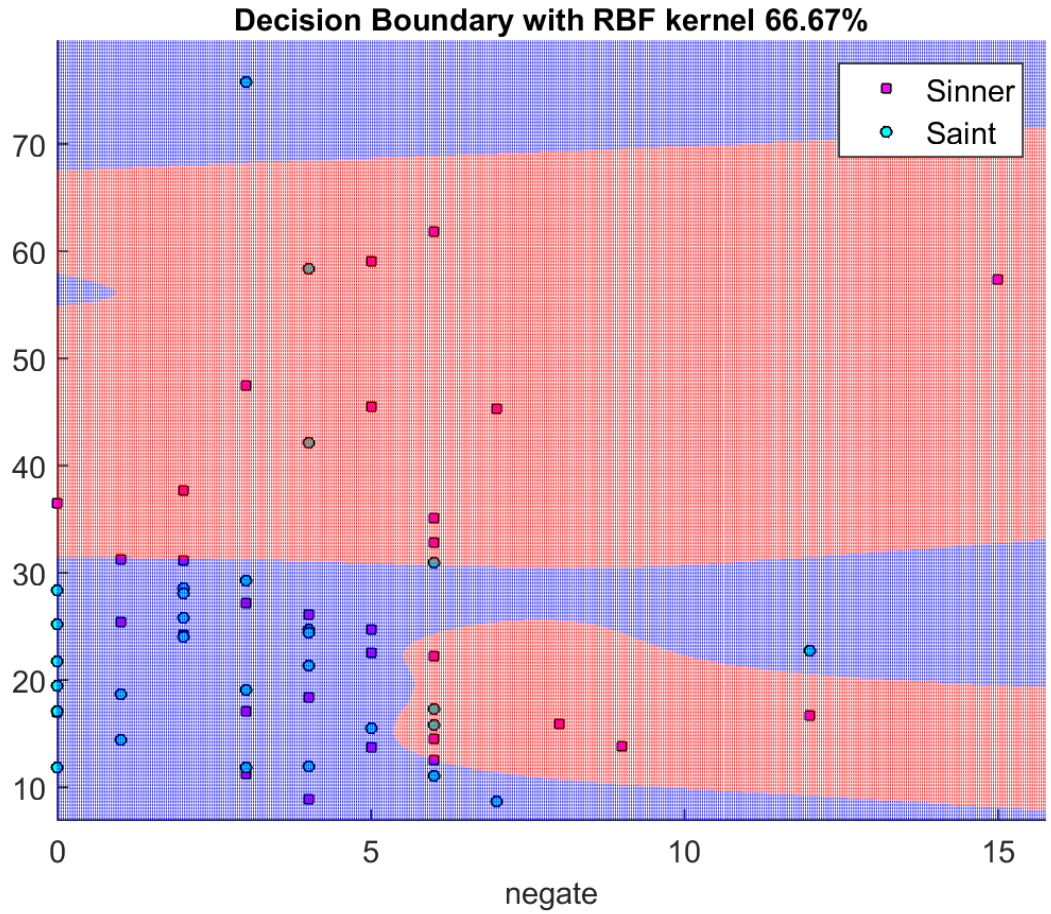
2D Support Vector Machine (SVM) RBF Kernel



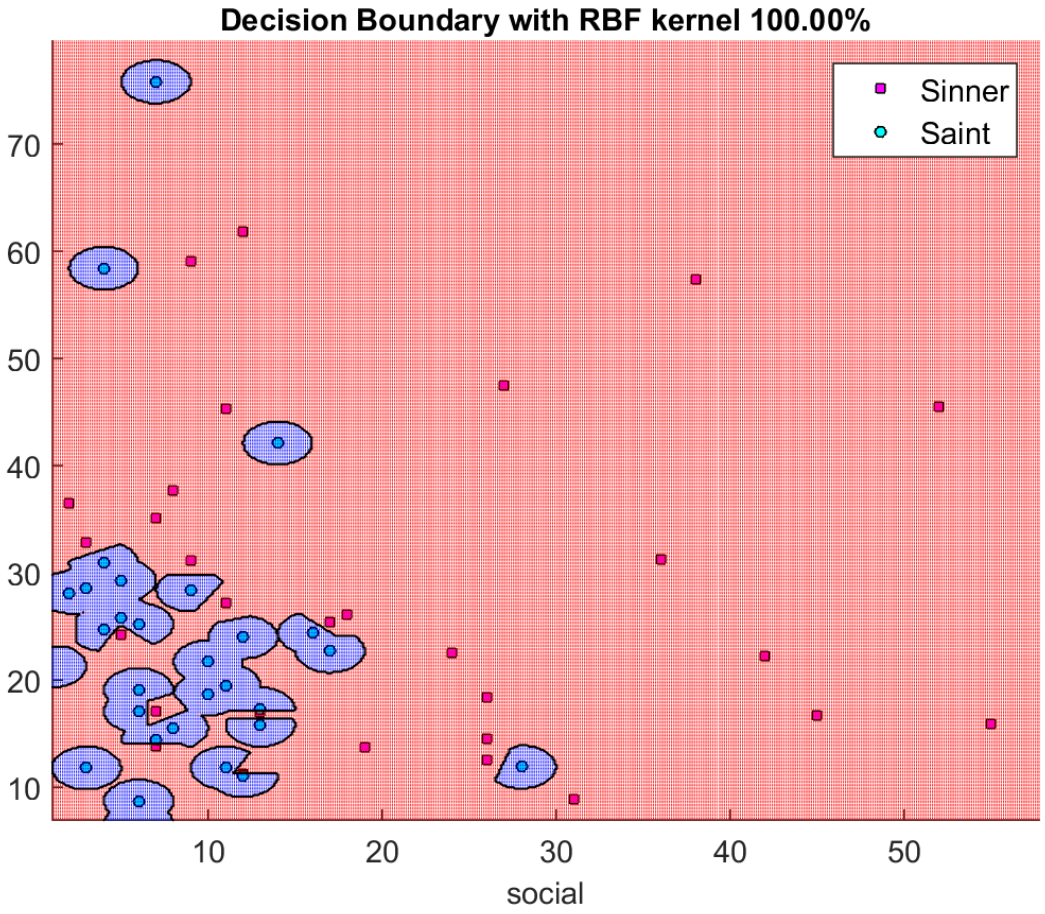


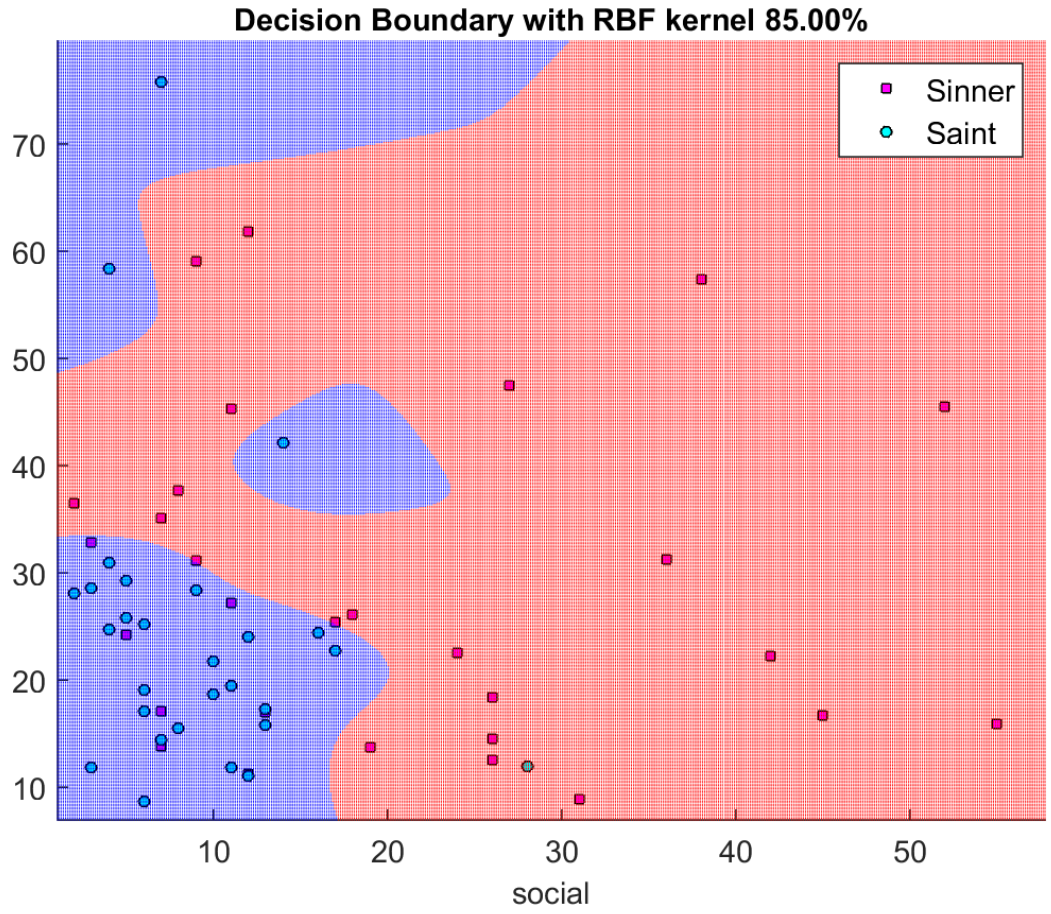
2D Support Vector Machine (SVM) RBF Kernel with leave-one-out cross-validation



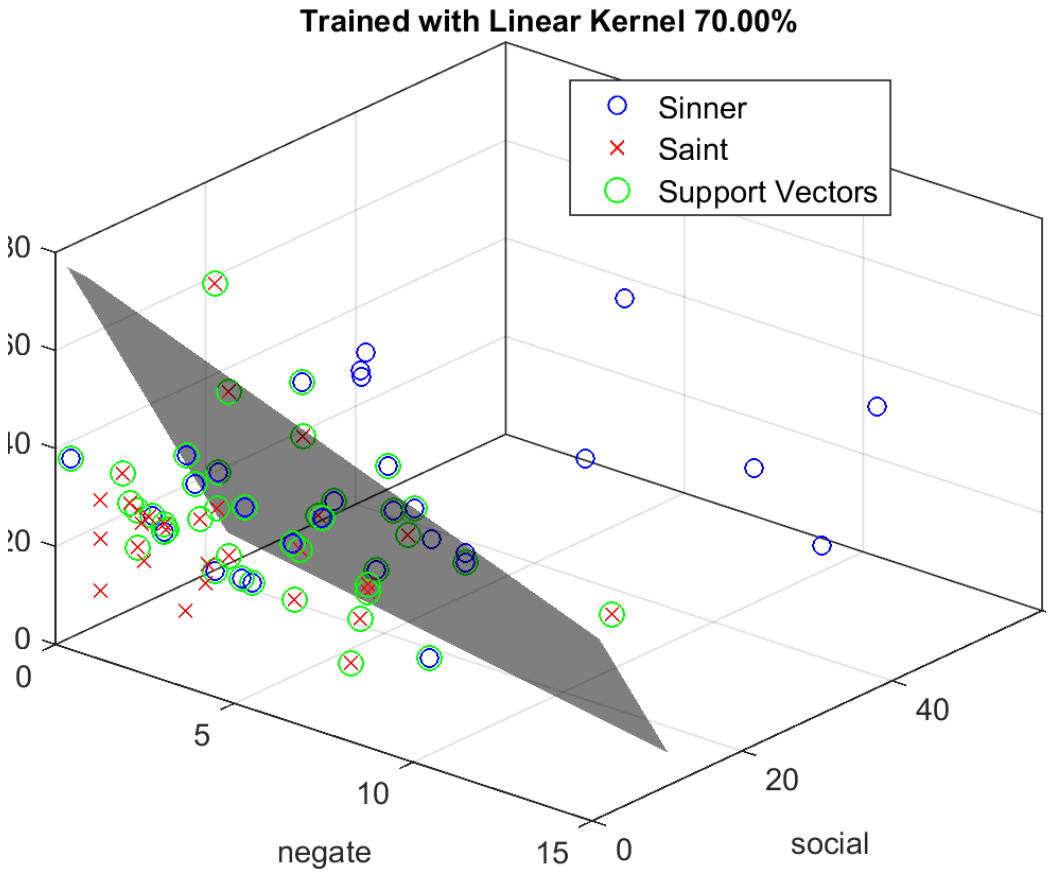


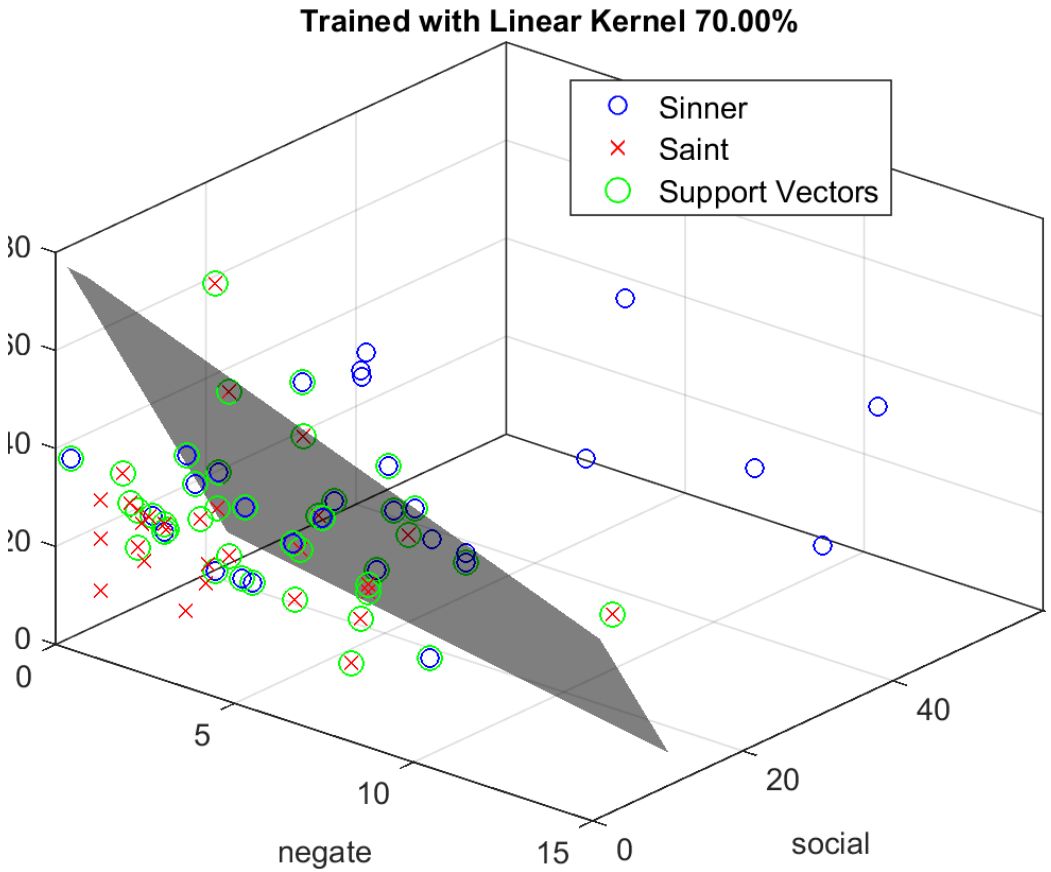
2D Support Vector Machine (SVM) RBF Kernel with leave-one-out cross-validation



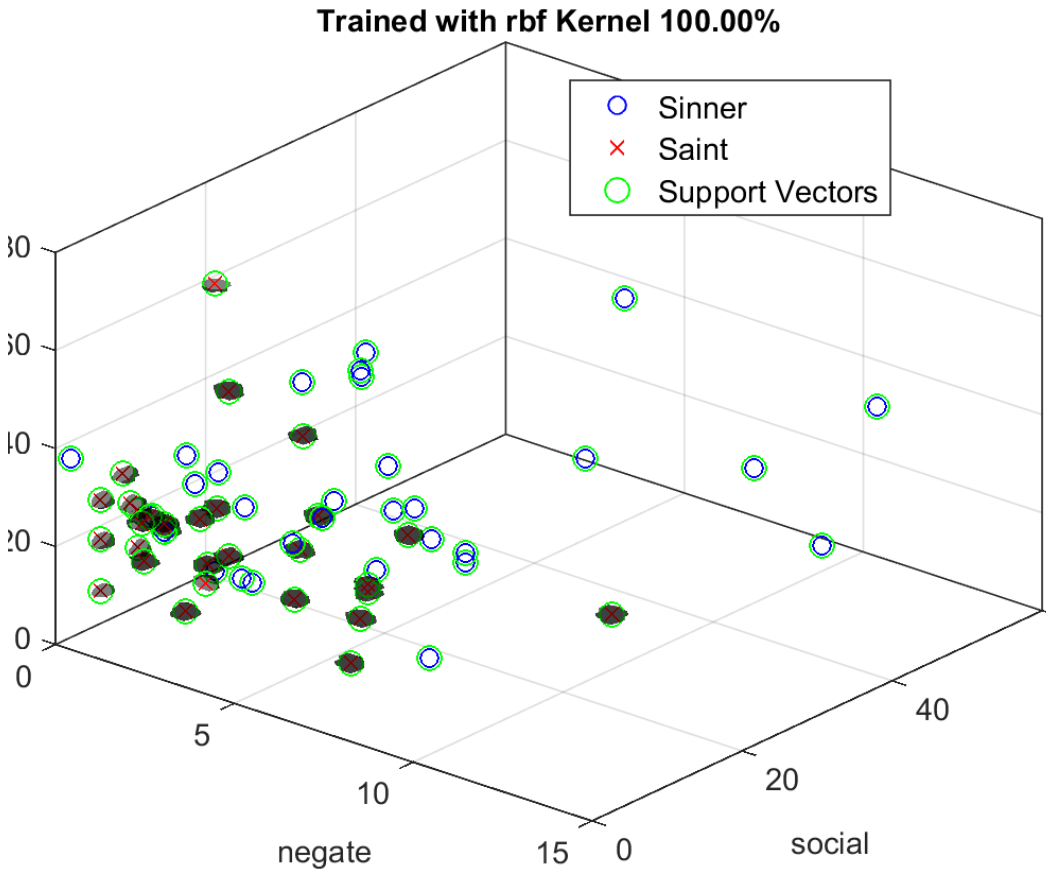


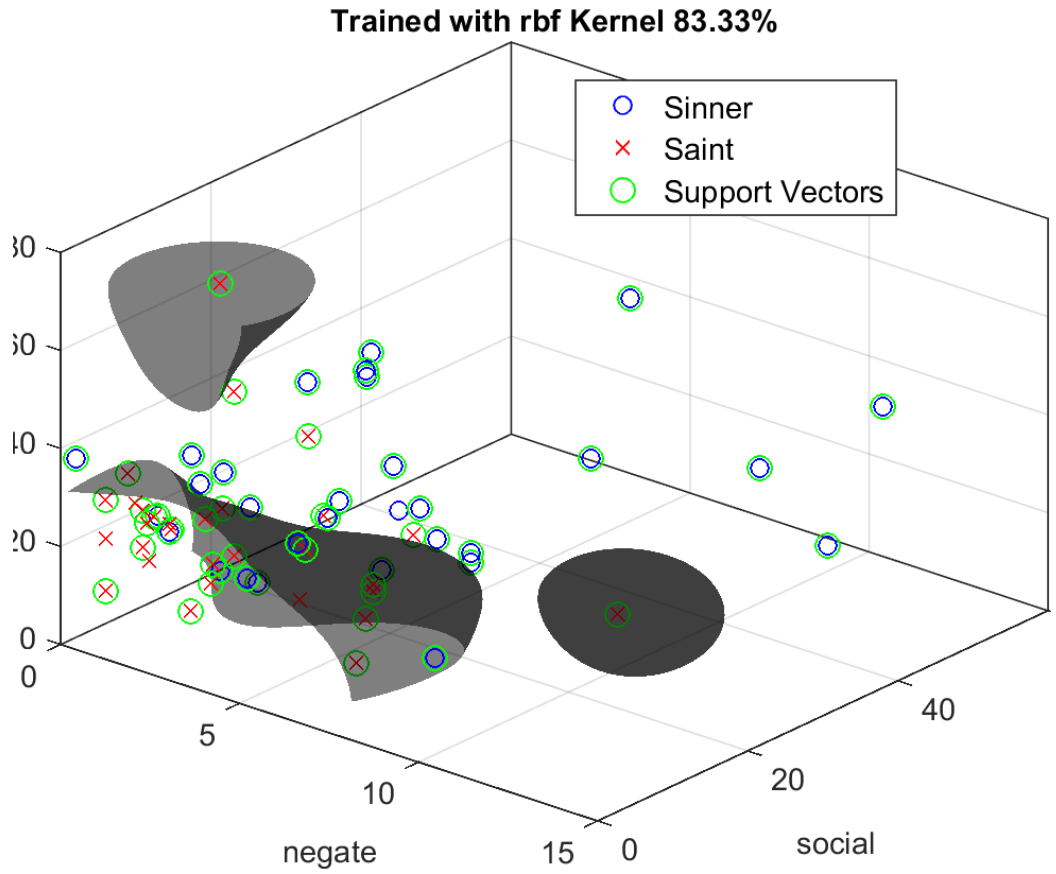
3D Support Vector Machine (SVM) Linear Kernel





3D Support Vector Machine (SVM) RBF Kernel with leave-one-out cross validation





Language-action Cues

Our findings suggest that certain language-action cues reveal patterns of information behavior manifested by deceivers in spontaneous online communication.

- ✓ *Cognitive load, Affective process*
- ✓ Negation, Latency (timelag), Social and Wordiness

Computational approaches to analyzing these language-action cues can provide significant accuracy in detecting computer-mediated deception.

LIMITATIONS AND CONCLUSION

LIMITATIONS

- This study was limited by small data/ sample size.
- Minor system and design issues were also a problem in each case.

CONCLUSIONS

- Online deception is a strategic process, but is also context-sensitive.
- Deception strategies differ based on context (i.e., cues can be different when attempting to identify a "simple" deceptive actor and when attempting to identify a deceptive actor in a group who may pose an insider threat), and modes (asynchronous vs. synchronous) of communication.
- By identifying key text-based cues in a data set, we can develop *predictive models* of human interactive behavior online and thus be in a position to *proactively* identify potential deceptive actors.

FUTURE RESEARCH

- Simulating additional deception scenarios with enhanced interactive and robust game designs
- **Online Polygraph**—automated detection systems of computer-mediated deception
- Capturing and studying vocal cues accompanying the text, using voice-to-text software.
- Design enhancements to both games are underway.
  - *Real or Spiel?* has been redesigned to use a platform developed in-house instead of Google+ Hangout.
  - The game interface itself has been redesigned
  - In future, *Real or Spiel?* will support multi-player (3+ players) interaction.

REFERENCES

- ✓Ho, S.M., Hancock, J.T., Booth, C., and Liu, X. (2016) Computer-mediated deception: Strategies revealed by language-action cues in spontaneous communication. **Journal of Management Information Systems**, Special issue on Systems Designed to Detect Deception, Fraud, Malicious Intent and Insider Threat. doi: 10.1080/07421222.2016.1205924.
- ✓Ho, S.M., Liu, X., Booth, C., and Hariharan, A. (2016) Saint or Sinner? Language-action cues for modeling deception using support vector machines. In K.S. Xu et al. (Eds.) **SBP-BRiMS**, LNCS 9708, 325-334. doi:10.1007/978-3-319-39931-7_31.
- ✓Ho, S.M., Hancock, J.T., Booth, C., Liu, X., Liu, M., Timmarajus, S.S. and Burmester, M. (2016) Real or Spiel? A decision tree approach for automated detection of deceptive language-action cues, **IEEE HICSS-49**, 3706-3715. IEEE: Kauai, Hawaii. doi:10.1109/ HICSS.2016.492.
- ✓Ho, S.M., Hancock, J.T., booth, C., Liu, X., Timmarajus, S.S., Burmester, M. (2015) Liar, Liar, IM on Fire: Deceptive language-action cues in spontaneous online communication. **IEEE Intelligence and Security Informatics (ISI)**, 157-159. IEEE: Baltimore, MD. doi: 10.1007/978-1-4799-9889-0/15.

Saint or Sinner? Language-action Cues for Modeling Deception Using Support Vector Machines

Shuyuan Mary Ho
smho@fsu.edu