

Applied Machine Learning Day - 2020

Facing Employers & Customers: What do Gaze & Expressions Tell About Soft Skills?

Skanda Muralidhar, Rémy Siegfried, Jean-Marc
Odobez, Daniel Gatica-Perez

About Me

- ➔ Post Doc @ Idiap Research Institute
- ➔ *Social Computing Group* headed by Prof Daniel Gatica-Perez
- ➔ integrate theories and models from ubiquitous computing, social media, machine learning, and social sciences, to sense, analyse, and interpret human and social behaviour in everyday life, and to create devices and systems that support interaction and communication.

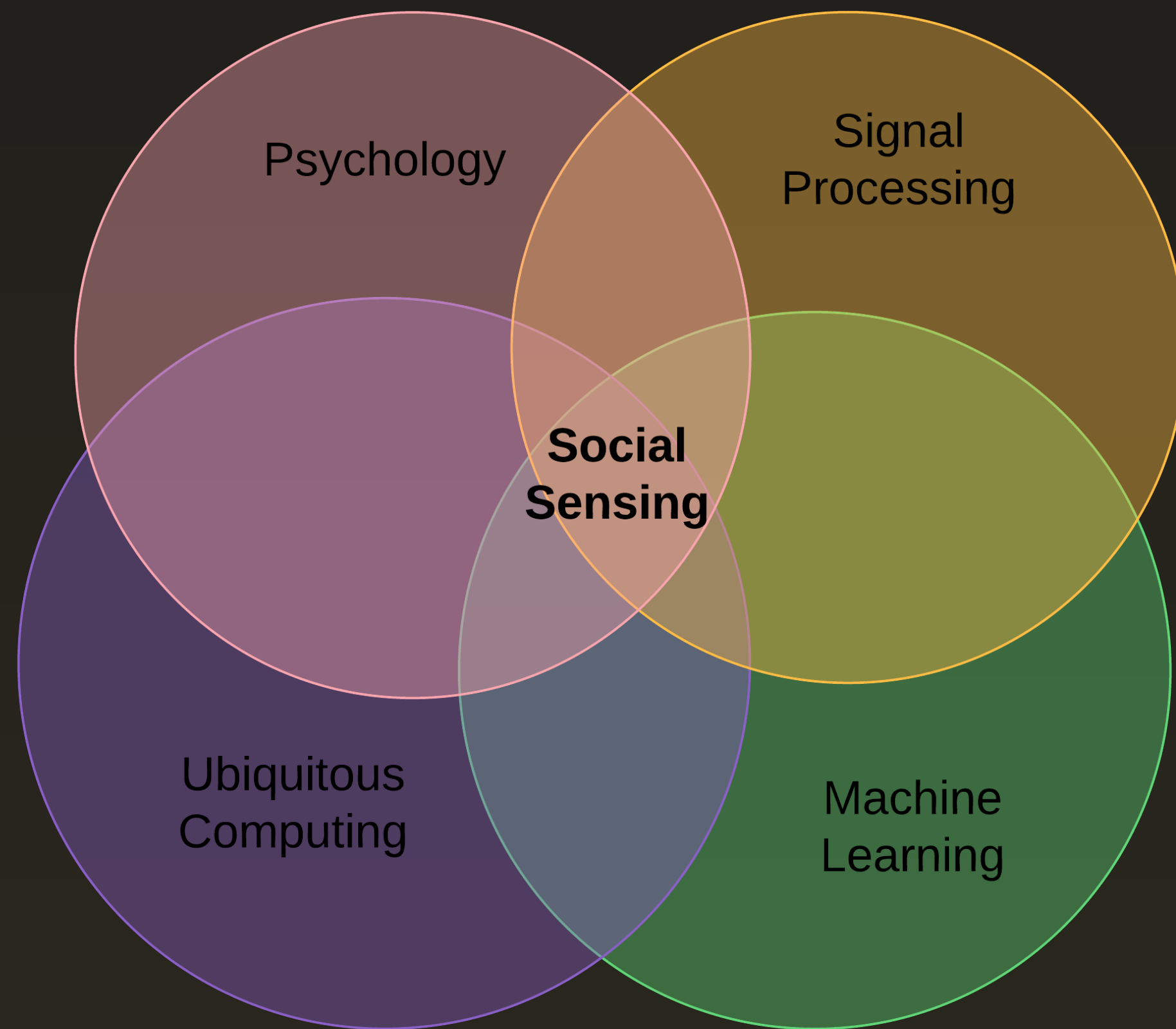


What am I doing?

Teach machines to predict automatically how individuals perceive one other and to model the cognitive processes that codify first impressions using **Social Sensing** approach.



What is Social Sensing



- ➔ Interaction between humans or between humans and computer consists of verbal and nonverbal behaviour.
- ➔ *Social Intelligence*: Ability to express and recognise social signals and behaviours.



Eye Gaze & Expressions

- Eye gaze & facial expressions - foundations of interpersonal communication.
- Contributes to formation of first impressions [Knapp, 2013]
- The mental image one forms about something or someone after a first encounter or meeting.



First Impressions Matter!

- Critical in hospitality industry [*Sundaram, 2000*]
- Customers form impression of organisations through service interactions
- Eye gaze - trust & credibility [*Beebe, 1980*]
[*Hemsley, 1978*]
- Facial expressions - interpersonal warmth [*Bayes, 1972*]



Literature in Computing

- ➔ Perceived Hirability [*Chen,2016*]
- ➔ Perceived Job Performance [*Muralidhar,2017*]
- ➔ Personality [*Batrinca, 2011*]
- ➔ Leadership [*Sanchez-Cortes, 2013*]

- ➔ So far, investigated in single workplace setting



Project Goals

- ➔ To develop a system for students to train themselves to improve their nonverbal behaviour
- ➔ UBImpressed dataset collected in collaboration with Vatel hospitality school, Martigny.
- ➔ Two important settings in hospitality industry
 - ➔ Employment interview
 - ➔ Hotel front desk
- ➔ Focus on interpretability not performance



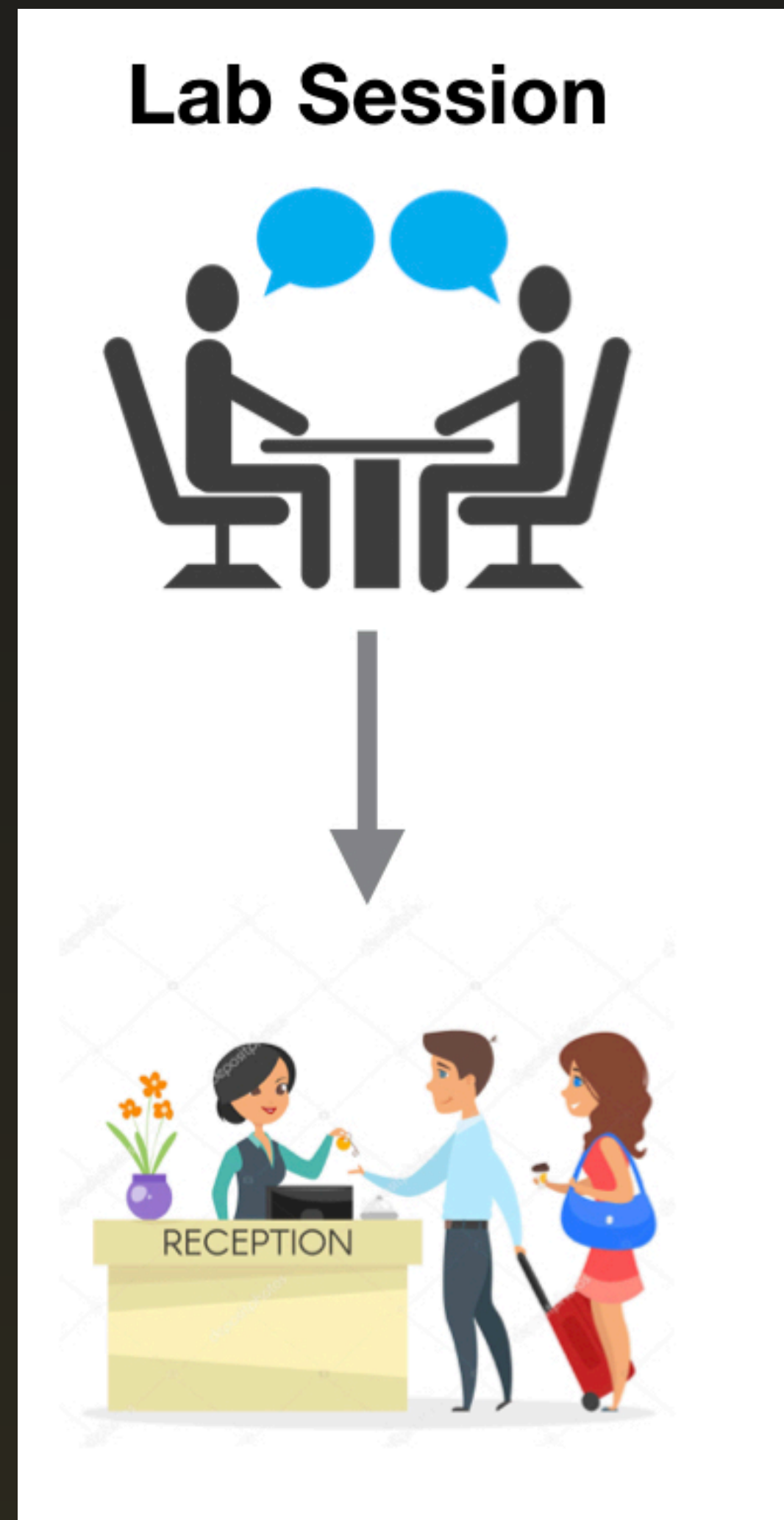
Objective

- ➔ Investigate connections between eye gaze, facial expressions
 - ➔ Perceived Hirability in practise job interviews
 - ➔ Job performance in practise reception desk interactions



Dataset

Data Corpus Collection¹



- 169 interactions in each situations (total 338 videos)
- Students of hospitality school
- Females - 57
- Males - 43
- Mean age - 20.6 yrs

¹Muralidhar et al 2016



Annotations





- ➔ Annotated by 5 Masters students
- ➔ Rated on Likert Scale of 1 - 7
- ➔ Inter Rater Agreement - ICC(2,k)
 - ➔ 0.52 - 0.73



- ➔ Annotated by 3 Masters students
- ➔ Rated on Likert Scale of 1 - 7
- ➔ Inter Rater Agreement - ICC(2,k)
 - ➔ 0.60 - 0.77

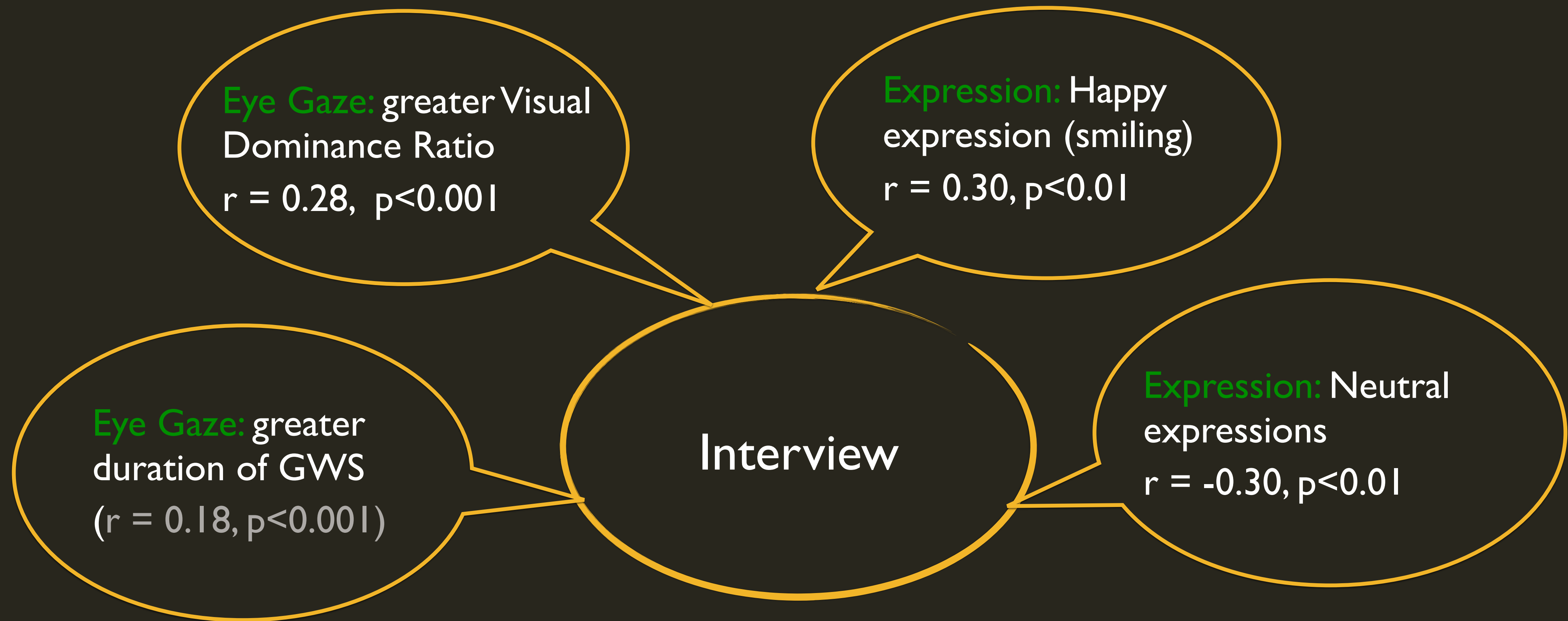


Features Extracted

Facial Expressions	Gaze [Siegfried 2017]	Visual	Audio
Sad	Gazing While Speaking (GWS)	Nodding while speaking	Speaking Ratio
Happy	Gazing While Listening (GWS)	Nodding while listening	Turn Duration
Surprised	Visual Dominance Ratio (VDR)	WMEI - Body Expressivity	Speaking Rate
Angry			Pauses
Disgust			Pitch
Fear			Speaking Energy
Neutral			Change in Speaking Energy
Contempt			

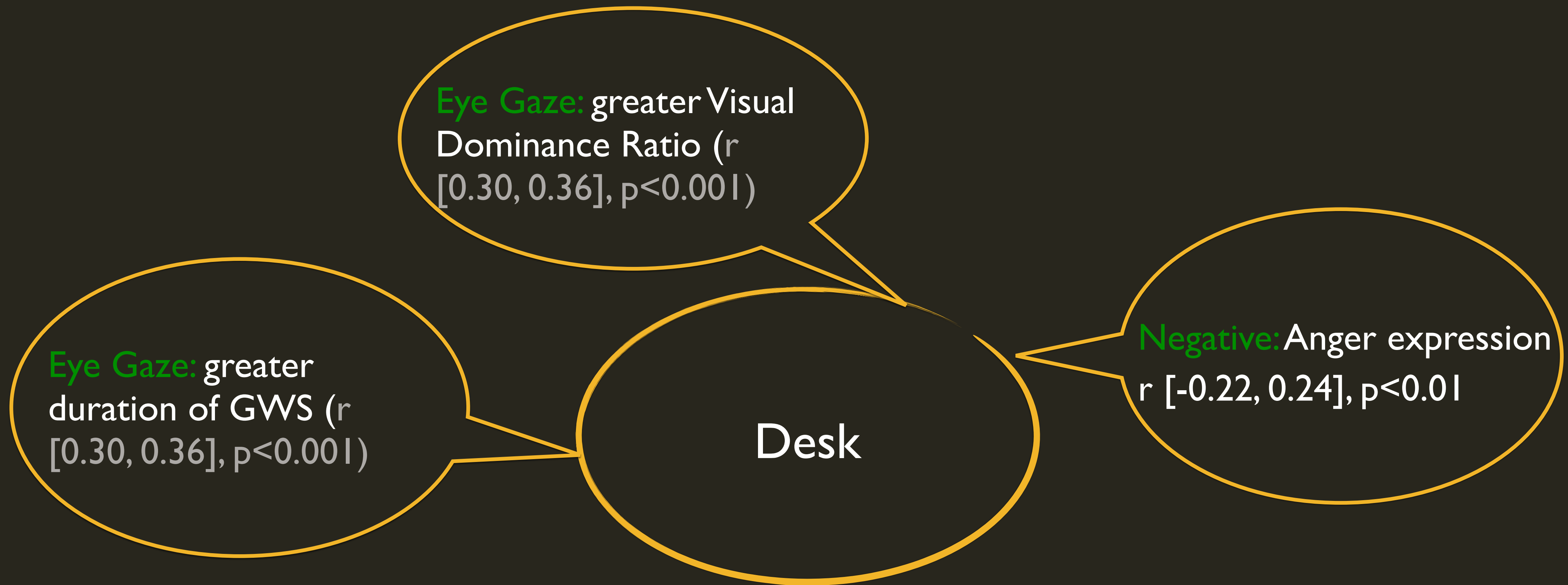


Correlation Analysis



➔ These results backed by literature [Amalfitano, 1977; Forbes & Jackson, 1980; Imada & Hakel, 1977]





➔ These results backed by literature [Soderlund & Rosengren, 2004; Sundaram & Webster 2000]



Inference

- ➔ Defined as regression task
- ➔ Evaluated using random forest (RF) algorithm
- ➔ Hyper-parameters tuned using 10-fold Cross-Validation (CV)
- ➔ Final scores obtained by Leave-one-video-out CV
- ➔ Repeated 100 times
- ➔ Evaluation Metric - Coefficient of determination (R^2)

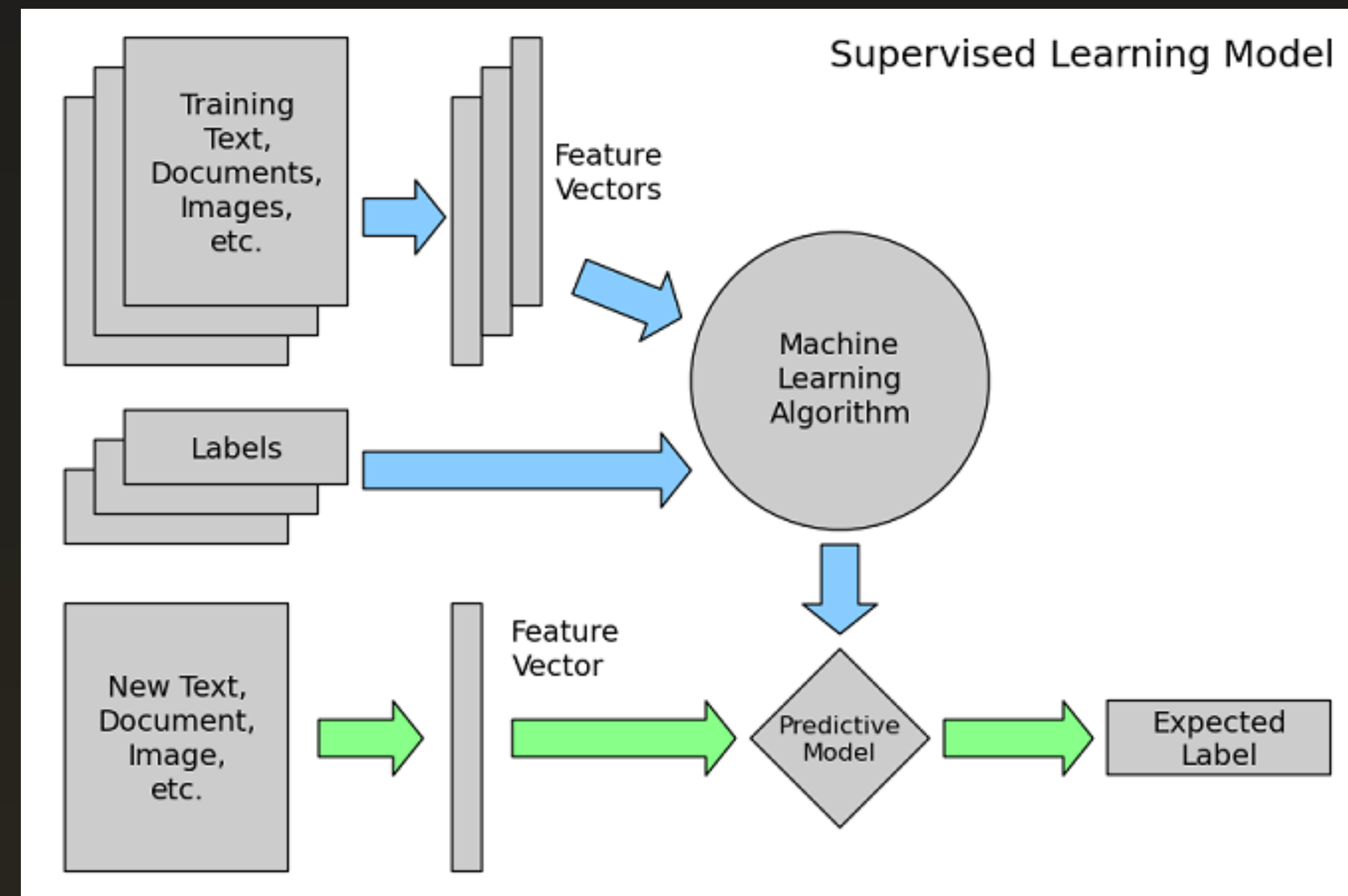
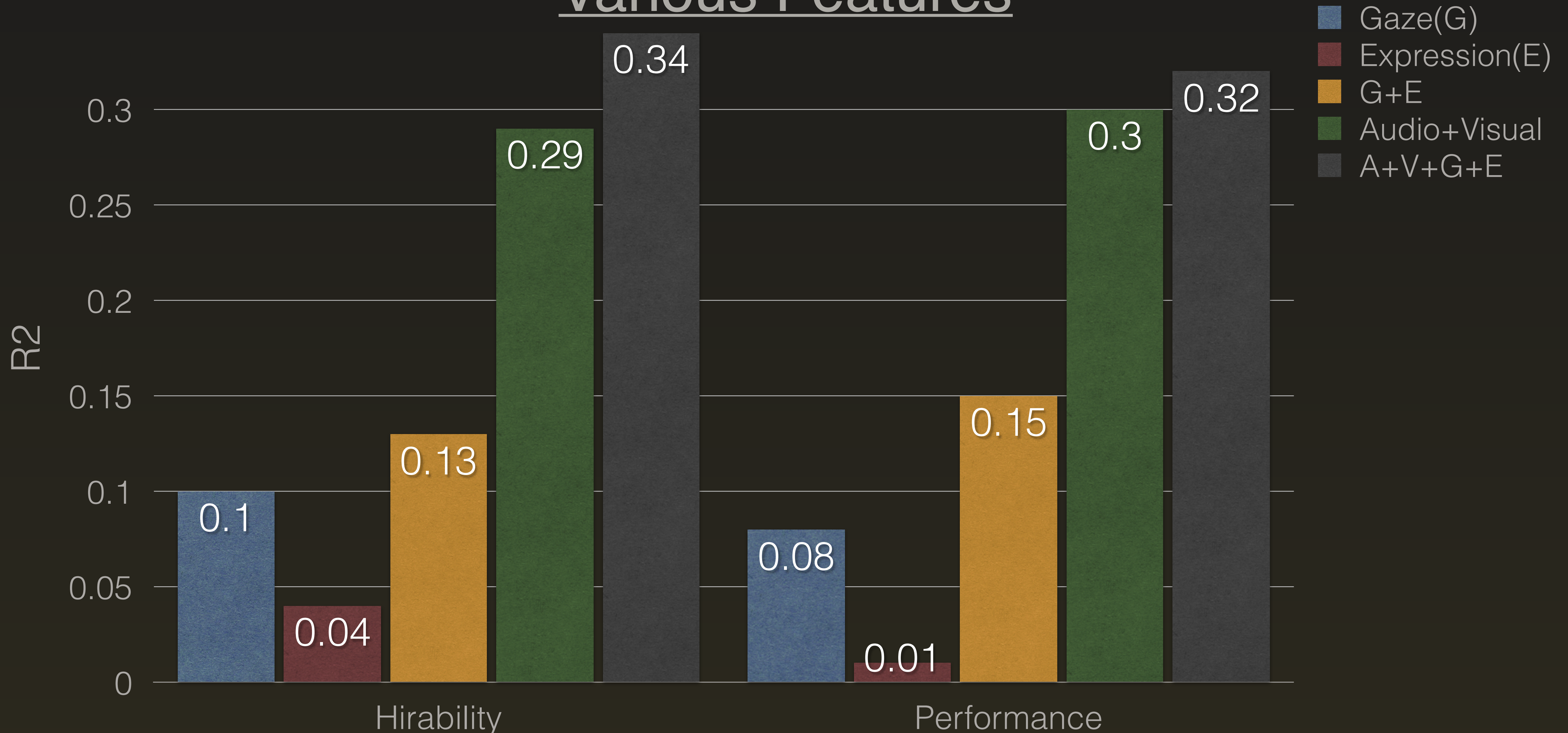


Image Source: <http://www.allprogrammingtutorials.com/tutorials/introduction-to-machine-learning.php>



Inferring Hirability & Performance from Various Features



Conclusions

- ➔ Moderate correlations between eye gaze, facial expressions and perceived soft skills in both settings.
- ➔ Low inference performance using eye gaze ($R^2 = 0.10$)
- ➔ Very low inference performance for facial expressions ($R^2 = 0.04$)
- ➔ Fusion of Gaze and Expressions with Audio-Visual provides the best inference performance ($R^2 = 0.34$)
- ➔ On going work investigating how feedback can be provided to individuals about their own nonverbal behaviour



ACKNOWLEDGMENTS

- ◆ Dr. Laurent Nguyen
- ◆ Dr. Denise Frauendorfer
- ◆ Research Assistants
- ◆ Prof. Marianne Schmid-Mast



Questions ?