Towards Improved Recognition and Diagnosis of Autism among Females – A Novel Approach using Machine Learning

Aanika Tangirala, Lakeside School, Seattle, WA, USA | BEHA023

BACKGROUND

Autism Spectrum Disorder (ASD) affects 1 in every 36 children[1]

Early intervention and **diagnosis** is key for successful outcomes in life[2]

For the **27% undiagnosed by age 8,** therapeutic intervention has dissipated[3]

Current diagnostic process is **time consuming, inaccessible, costly,** and **subjective**

It can take up to **13 months** to get a diagnosis

ASD is diagnosed solely on **behavioral observation** – no reliable biomarkers

THE HIDDEN ISSUE: GENDER BIAS IN DIAGNOSIS

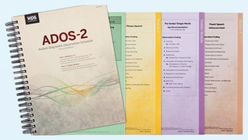
- Females with autism are often misdiagnosed or miss diagnosis due to biased diagnostic tests[4]
- Leading Theories
  - **Female Autistic Phenotype** exists[5]
  - Females tend to **camouflage** and **mask** symptoms[5]
- Delayed diagnoses and support leads to higher % of **anxiety** and **depression** in teen girls with autism compared to male peers[6]

Difference cannot ONLY be attributed to something inherent in females that decreases the chances of developing autism.

CURRENT AUTISM DIAGNOSTIC TESTS

The **Diagnostic and Statistic Manual of Mental Disorders, Fifth Edition (DSM-5)** provides criteria for diagnosing ASD based on impairments in social and verbal communication and restricted, repetitive patterns of behavior

- **Autism Diagnostic Observation Schedule (ADOS)** – Clinical Observation
  - ADOS_COMM: communication ability
  - ADOS_SOCIAL: social ability
  - ADOS_STEREO_BEHAV: repetitive behavior
- **Autism Diagnostic Interview-Revised (ADI-R)** – Caregiver Interview
  - ADI_R_SOCIAL: social ability
  - ADI_R_VERBAL: verbal ability
  - ADI_R_RRB repetitive & restrictive behavior:
- **Social Communication Questionnaire (SCQ)** – Caregiver Survey
  - SCQ_TOTAL: social and communication ability
- **Social Responsive Scale (SRS)** – Caregiver Survey
  - SRS_AWARENESS
  - SRS_COGNITION
  - SRS_COMMUNICATION
  - SRS_MOTIVATION
  - SRS_MANNERISMS



Autism Diagnostic Observation Schedule (ADOS-2) 2nd Edition
Pearson Clinical

Sample SCQ Form
Children's Hospital of Philadelphia Research Institute

Sample SRS-2 Form
Mind Resources

Clinicians use a combination of above assessments to diagnose autism

RESEARCH QUESTION AND GOALS

Research Question: Which **features** of existing Autism Spectrum Disorder (ASD) diagnostic tests are most useful for **predicting autism** in **female** populations?

1) Establish **diagnostic gender bias** and behavioral differences between females and males with autism in diagnostic tests.

2) Identify **minimal** set of behaviors and features most predictive of autism in females.

3) Create a **more accurate diagnostic predictive** machine learning model for females.

DATA AND METHODS

DATA SOURCE

- **ABIDE** – the Autism Brain Imaging Data Exchange[7]
- Public online repository of phenotypic autism behavioral data collected over **17** global sites
- Contains **73** features across ADOS, ADI-R, SRS, SCQ
- Dataset size = **1112 individuals, 948 males, 164 females**
- **539** individuals diagnosed with autism and **573** without

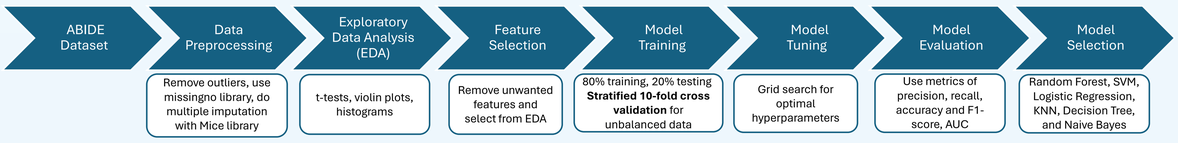
DATA PREPROCESSING

- Remove outliers
- Remove features not related to ADOS, ADI-R, SCQ, SRS tests, from 73 to 23 features
- Use multiple imputation methods for missing data based on correlation heatmap

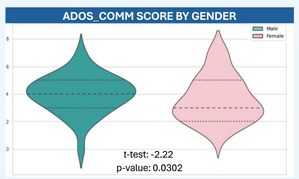
EXPLORATORY DATA ANALYSIS (EDA)

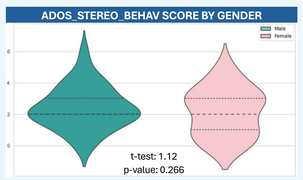
- Descriptive statistics of the features
- Visualization of diagnostic features by gender using histograms and violin plots
- Check for candidate gender-specific features for feature selection and towards Goal 1

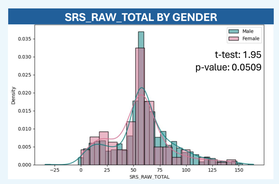
MACHINE LEARNING PIPELINE

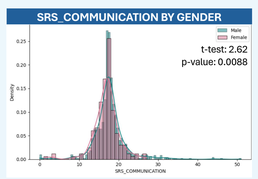


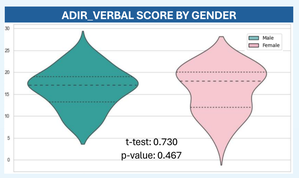
DATA ANALYSIS AND RESULTS

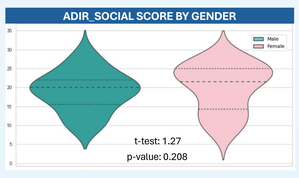


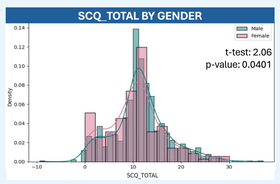












- Males have much sharper distributions centered around the means
- Females have bimodal distributions

- Females have higher scores in ADI_R tests
- Statistical difference in ADOS_COMM between males and females

Statistical significance ($p \leq 0.05$) between genders identified with t-tests in questionnaire-based features – **SRS_RAW_TOTAL, SRS_COMMUNICATION,** and **SCQ_TOTAL**: potential candidates for **feature selection** for female model

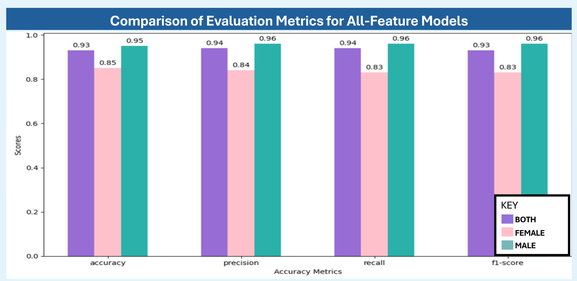
Diagnostic tests capture and classify female autistic behavior differently than males

1) ESTABLISHED DIAGNOSTIC GENDER BIAS AND BEHAVIORAL DIFFERENCES IN CURRENT TESTS

(1) Built Random Forest classifier on processed ABIDE dataset with all 23 features to predict autism

(2) Trained and tested on male and female data combined **(Model 1)**, on female dataset only **(Model 2)**, on male dataset only **(Model 3)**

(3) Checked model evaluation metrics: accuracy, precision, recall, etc.



Model	Accuracy	Precision	Recall	F1 score
1:Both	0.93	0.94	0.94	0.93
2: Females	0.85	0.84	0.83	0.83
3: Males	0.95	0.96	0.96	0.96

Female dataset has least accuracy on all-feature model, Males have most accuracy
Validates gender bias of tests

2) IDENTIFIED MINIMAL FEATURE SET FOR AUTISM PREDICTION IN FEMALES

(4) Checked Top 10 most impactful features for each model. Gender based differences observed

(5) Using Top 10 + candidates from EDA, built minimal feature set of **12 features**

Comparison of Top 10 Impactful Features Across Genders

Top 10 features for males + females	Top 10 features for females	Top 10 features for males
AQ_TOTAL	ADI_R_SOCIAL_TOTAL	AQ_TOTAL
ADI_R_VERBAL_TOTAL	ADI_R_VERBAL_TOTAL	ADOS_GOTHAM_SEVERITY
ADOS_SOCIAL	SCQ_TOTAL	ADOS_SOCIAL
ADI_RRB_TOTAL	ADOS_COMM	ADI_RRB_TOTAL
SCQ_TOTAL	SRS_RAW_TOTAL	ADI_R_VERBAL_TOTAL
ADOS_COMM	AQ_TOTAL	ADOS_COMM
ADOS_GOTHAM_SEVERITY	ADOS_SOCIAL	ADOS_TOTAL
ADOS_TOTAL	ADOS_GOTHAM_RRB	SRS_RAW_TOTAL
ADI_R_SOCIAL_TOTAL	ADOS_GOTHAM_SOCAFFECT	SCQ_TOTAL
SRS_RAW_TOTAL	ADOS_GOTHAM_SEVERITY	SRS_COGNITION

ADOS_TOTAL, ADIR_RRB not **impactful** for prediction in female dataset

Interview (ADI_R) and **Questionnaire (SCQ, SRS)** features more impactful for females

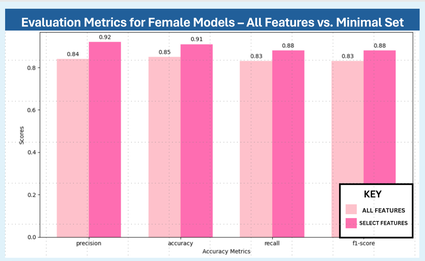
Minimal Set of 12 most predictive features for females established

SCQ TOTAL	ADOS COMM
SRS RAW TOTAL	AQ TOTAL
SRS AWARENESS	ADOS SOCIAL
SRS COMMUNICATION	ADOS GOTHAM RRB
ADIR SOCIAL TOTAL	ADOS GOTHAM SOCAFFECT
ADIR VERBAL TOTAL	ADOS GOTHAM SEVERITY

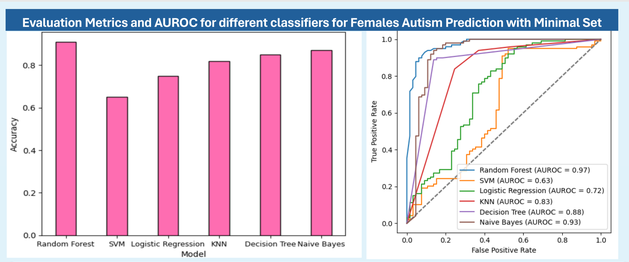
3) CREATED MORE ACCURATE AUTISM DIAGNOSTIC PREDICTIVE ML MODEL FOR FEMALES

(6) With minimal set of 12 features, built 6 different classifiers for females

(7) Compared evaluation metrics of model with All Features to model with Minimal Set Features



Minimal Set feature model for females more accurate than All Features model (91% vs. 85%)



Random Forest is most accurate classifier and has largest AUROC on Minimal Features Set model for Females

CONCLUSIONS

- Significant differences in behavior of males and females with autism and their performance on social, communication, and repetitive subtests; **supports Female Autistic Phenotype hypothesis**
- **Gender bias in current diagnostic tests proven**, more accurate in diagnosing males than females (95% vs 85%)
- **Random Forest classifier** with **minimal set of 12 features** (from original 23, **48% reduction**) was more accurate in predicting female autism **(91% vs 85%)**
- **Questionnaire and caregiver interview-based tests** were better indicators for female autism; supports camouflaging and masking behavior in females.
- More female data across multiple sites and access to **item-specific scores** would potentially lead to even deeper insights

FUTURE RESEARCH

- Working with Children's National Neurodevelopmental Program (Washington DC) for deeper item level sub-score analysis
- Similar analysis and models for removing bias based on race and socioeconomic factors
- Generative adversarial networks (GANs) to address lack of female data issues
- Combination of behavioral and neuroimaging data to provide more accurate diagnosis

APPLICATIONS OF RESEARCH

- Modify DSM-5 and current diagnostic tests for females - **remove overreliance** on ADOS test and repetitive behaviors sub-tests
- Weight **questionnaire and interview-based** responses more heavily; use these for early screening
- Develop **digital and mobile health approaches** for faster and expedient screening
- **Raise awareness** for caregivers, teachers, and parents to **identify female autistic behaviors** and flag for earlier diagnosis

REFERENCES

- [1]https://www.cdc.gov/ncbddd/autism/data.html
- [2]Okoye, Chiugo et al. "Early Diagnosis of Autism Spectrum Disorder: A Review and Analysis of the Risks and Benefits." Cureus vol. 15,8 e43226. 9 Aug. 2023, doi:10.7759/cureus.43226
- [3]Shattuck PT, Durkin M, Maenner M, Newschaffer C, Mandell DS, Wiggins L, et al. Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. J Am Acad Child Adolesc Psychiatry. 2009;48:474–483.
- [4]Werling, Donna M, and Daniel H Geschwind. "Sex differences in autism spectrum disorders." Current opinion in neurology vol. 26,2 (2013): 146-53. doi:10.1097/WCO.0b013e32835ee548
- [5]Hull, L., Petrides, K.V. & Mandy, W. The Female Autism Phenotype and Camouflaging: a Narrative Review. Rev J Autism Dev Disord 7, 306–317 (2020). https://doi.org/10.1007/s40489-020-00197-9
- [6]Hartley SL, Sikora DM. Sex differences in autism spectrum disorder: an examination of developmental functioning, autistic symptoms, and coexisting behavior problems in toddlers. J Autism Dev Disord. 2009 Dec;39(12):1715-22. doi: 10.1007/s10803-009-0810-8. Epub 2009 Jul 7. PMID: 19582563; PMCID: PMC3590797.
- [7]https://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html
- Kosmicki JA, Sochat V, Duda M, Wall DP. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Transl Psychiatry. 2015 Feb 24;5(2):e514. doi: 10.1038/tp.2015.7. PMID: 25710120; PMCID: PMC4445756.
- Rahman MM, Usman OL, Muniyandi RC, Sahran S, Mohamed S, Razak RA. A Review of Machine Learning Methods of Feature Selection and Classification for Autism Spectrum Disorder. Brain Sci. 2020 Dec 7;10(12):949. doi: 10.3390/brainsci10120949. PMID: 33297436; PMCID: PMC7762227.
- Hirokazu Kumazaki, Taro Muramatsu, Hirotaka Kosaka, Takashi X. Fujisawa, Kazuhiko Iwata, Akemi Tomoda, Kenji Tsuchiya, Masaru Mimura, Sex differences in cognitive and symptom profiles in children with high functioning autism spectrum disorders, Research in Autism Spectrum Disorders, Volumes 13–14, 2015, Pages 1-7, ISSN 1750-9467, https://doi.org/10.1016/j.rasd.2014.12.011.
- Schuck, R. K., Flores, R. E., & Fung, L. K. Sex/Gender Differences in Symptomology and Camouflaging in Adults with Autism Spectrum Disorder. Journal of autism and developmental disorders, (2019) 49(6), 2597–2604. https://doi.org/10.1007/s10803-019-03998-y

All figures are student created except where cited