Predictive Modeling and Explainable AI for Autism Spectrum Disorder Diagnosis

Advisor: Dr. Amir Jafari and Dr. Gabriela Rosenblau
Students: Anirudh Rao, Ramana Bhaskar Kosuru, Wilona Nguyen

Problem Statement

How can interpretable, multi-modal machine learning models that integrate behavioral and linguistic data improve the diagnostic precision, subtype discovery, and generalizability of Autism Spectrum Disorder (ASD) prediction?

Objective

Extract insights from unstructured text data

↓

Implement feature importance analysis

↓

Improve autism diagnosis accuracy

↓

Create generalize models

↓

Identify autism subclusters

↓

Develop explainable AI framework

Raw Data Description - Data V1

- Provided by Department of Psychological & Brain Sciences
- 1119 subjects (ages 8-12Y)

sub	Participant ID
td_or_asd	Is this participant classified as TD (Typically Developing) or ASD (Autism Spectrum Disorder)? 0 = TD, 1 = ASD
profile	TDprof_norm, ASDprof_norm, or ASDprof_unif. These represent the preference profiles of the 3 kids that participants learned about. Participants will have data for at least 2 profiles.
subject	Participant ID followed by a random code for the current profile
asd_diagnosis_text	Which ASD diagnosis does the participant have? Only available for ASD participants. (Aspergers Disorder, Autism, Child Disintegrative Disorder (CDD), PDD-NOS, Other or Unknown, NA)
trial	The trial number (1-60) within the current profile.
image	The image shown on the current trial. Order is randomized within profile for each participant.
cat	Image category (1 = activities, 3 = foods)
subcat	Image subcategory (1 = arts & crafts, 2 = music & instruments, 3 = sports, 4 = games & gadgets, 9 = fast food, 10 = fruits & vegetables, 11 = healthy & savory, 12 = desserts)

Raw Data Description - Data V1

concept	Descriptive image concept label (eg. asian_cuisine, baked_good, water_sport, instrument,....)
cnum	Image concept number (1-25)
selfpref	How much does the participant like this image? (0-100, NA)
asd_meanpref	On average, how much do ASD participants like this image?
td_meanpref	On average, how much do TD participants like this image?
slider_rating	The participant's rating of how much they think the peer in question like the image shown. (1-21)
profile_rating	The correct rating for the given image and profile (this is the feedback given to participants) (1-21)
PE	Prediction Error. The absolute difference between a participant's rating of how much they think the peer likes an item and the peer's actual liking of the item (actual liking = feedback).
profile_avg_PE	Participant's average prediction error (PE) on each peer profile. Prediction error was calculated as the absolute difference between a participant's rating of how much they think the peer likes an item (slider_rating) and the peer's actual liking of the item (profile_rating).
free_response	Participant's free response -- what they think about the peer they just learned about.

Raw Data Description - Data V1

- LLM Trial Level Data .csv
- 1119 subjects (ages 8-12Y), 3 profiles (TDprof_norm, ASDprof_norm, or ASDprof_unif.)
- Each subject has 60 trials for each profile
- **187,187** observations and **19** variables

sub	subject	td_or_asd	asd_diagnc	trial	profile	image	cat	subcat	concept	cnum	selfpref	asd_meanp	td_meanpr	slider_ratin	profile_rati	PE	profile_avg	free_response
3HLdZX	3HLdZX_0p	1	Aspergers [	1	ASDprof_nc	bagel_1	3	11	bagel	20	NA	57.18412	56.68449	16	15	1	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	2	ASDprof_nc	chess_2	1	4	games	12	87	53.71631	62.63415	11	9	2	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	3	ASDprof_nc	cookie4	3	12	baked_goo	24	100	84.58782	85	21	10	11	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	4	ASDprof_nc	beads_1	1	1	crafting_su	2	NA	50.18051	36.8984	14	11	3	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	5	ASDprof_nc	pencil_3	1	1	art_supplie	1	NA	69.74729	61.06952	10	5	5	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	6	ASDprof_nc	notebook_1	1	1	writing_sup	3	NA	48.15884	40	12	17	5	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	7	ASDprof_nc	bun_1	3	12	baked_goo	24	NA	75.01805	67.80749	16	8	8	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	8	ASDprof_nc	ice_skates_	1	3	skating	9	NA	34.29603	33.15508	8	1	7	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	9	ASDprof_nc	domino	1	4	games	12	NA	51.98556	47.37968	9	7	2	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	10	ASDprof_nc	asian_01	3	11	asian_cuisi	18	56	38.42776	50.13659	11	14	3	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	11	ASDprof_nc	dice_1	1	4	games	12	58	54.22238	55.05172	14	17	3	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	12	ASDprof_nc	mikado_2	1	4	games	12	NA	41.44404	42.03209	8	14	6	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	13	ASDprof_nc	skateboard	1	3	skating	9	NA	43.2491	48.6631	14	10	4	4.9	They like sports but not Bolling he enj
3HLdZX	3HLdZX_0p	1	Aspergers [	14	ASDprof_nc	cam_corde	1	4	photograph	13	34	58.89518	61.48662	14	4	10	4.9	They like sports but not Bolling he enj

Data V2 Description - Aggregated Level (sub | profile)

- LLM Data.csv
- 1119 subjects (ages 8-12Y), 3 profiles (TDprof_norm, ASDprof_norm, or ASDprof_unif.)
- Each subject has 3 profiles (avg_PE is calculated across 60 trials for each profile)
- **2,648** observations and **11** variables

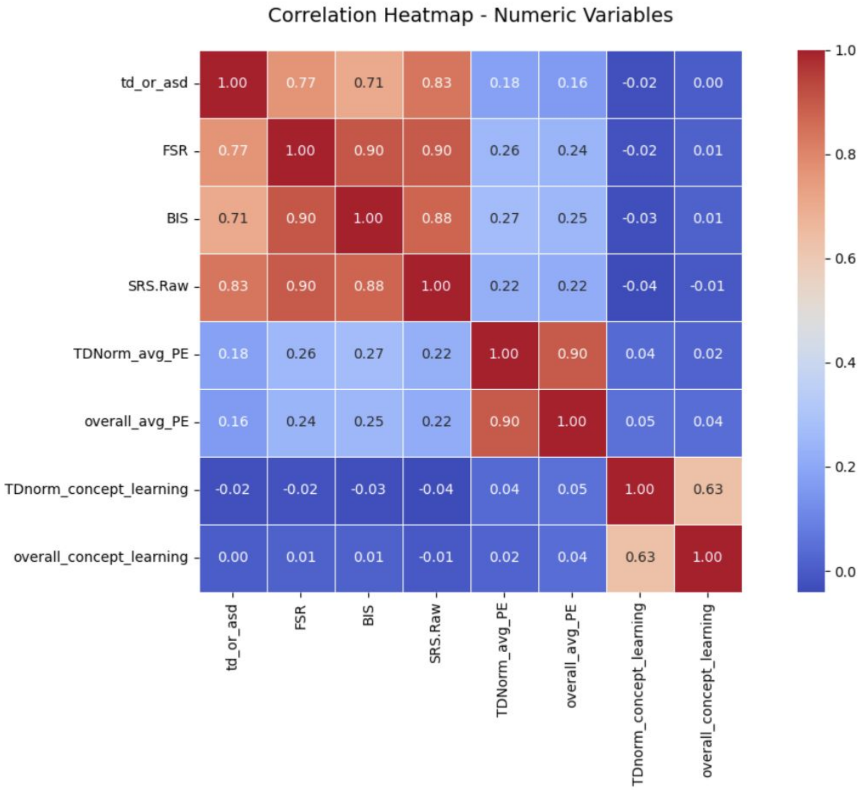
sub	profile	subject	td_or_asd	SRS.Raw	FSR	BIS	avg_PE	free_respons	LPA_Profile_g	LPA_Profile_ASD_only
3HLdZX	ASDprof_norı	3HLdZX_0p0	1	NA	20	25	4.9	They like spoı	1	1
3HLdZX	ASDprof_unif	3HLdZX_qAFI	1	NA	20	25	7.01666667	all i can say fı	1	1
3HLdZX	TDprof_norm	3HLdZX_z2C	1	NA	20	25	4.4	They enjoy al	1	1
4EZM27	ASDprof_norı	4EZM27_lEZI	1	103	54	100	6.23333333	he likes taco	1	1
4EZM27	ASDprof_unif	4EZM27_B36	1	103	54	100	6.68333333	I learned tha	1	1
4EZM27	TDprof_norm	4EZM27_kcn	1	103	54	100	4.66666667	They like mus	1	1
4ZEZ5S	ASDprof_norı	4ZEZ5S_xAW	1	126	52	161	6.76666667	Rowan is 12y	2	4
4ZEZ5S	ASDprof_unif	4ZEZ5S_709	1	126	52	161	8.65	Quinn does n	2	4
4ZEZ5S	TDprof_norm	4ZEZ5S_qc6	1	126	52	161	4.6	Parker does r	2	4

Data V3 Description - Participant Level (sub)

- LLM data_aggregate_8.25.25 data_updated 10.8.25.csv
- 1119 subjects (ages 8-12Y), 3 profiles (TDprof_norm, ASDprof_norm, or ASDprof_unif.)
- Each subject has:
  - **TDNorm_avg_PE**: average PE for TDNorm profile
  - **overall_avg_PE**: average PE across all 3 profiles
  - **TDNorm_concept_learning**: slope of PE for TDNorm profile
  - **overall_concept_learning**: slope of PE across all 3 profiles
- **1,119** observations and **10** variables

sub	td_or_asd	FSR	BIS	SRS.Raw	TDNorm_avg_PE	overall_avg_PE	TDnorm_concept_learning	overall_concept_learning	free_response_TDprof_norm
235X4W	0	5	4	10	5.716666667	7.111111111	-0.243313315	-0.004524569	He likes ice cream, kind of likes chess.
26NE6D	0	15	6	11	3.95	5.561111111	0.09796008	0.021768026	he doesn't like dice, skating, and exercising. really likes ch
27YMQK	0	NA	2.054054054	12	7.766666667	7.95	-0.021473482	0.020730875	he was normal
2A7DFT	0	7	0	7	4.266666667	4.938888889	-0.127769397	-0.151312665	She likes ice skating, fruit, the accordion, tasty food. She
2BMMAM	0	24	34	19	6.416666667	6.491666667	-0.339128115	-0.208756422	Quinn like's basketball she does not like game's she does
2G3JNL	0	11	9	20	4.633333333	5.822222222	0.060038511	0.012699918	Not a very pickey eater. Dislikes workouts. Detests most d
2GJdWc	0	3	12	18	4.566666667	5.644444444	0.138521821	0.069401618	Alex most notably liked more sophisticated foods, such a
2HSJ5Z	0	8	0	23	3.9	5.161111111	0.024022213	0.167810196	Alex likes food, especially fruits and sweets. He also really
2UAdZU	0	18	6	15	4.35	4.927777778	0.193282662	0.03618996	likes to eat pretty healthy
2UK6c5	0	17	18	20	5.183333333	5.755555556	0.246659019	0.069541374	He likes most musical instruments. He likes some ice ska
2WNcEH	0	41	48	58	5.5	6.388888889	-0.107478777	-0.14833702	he ok to plqy piano. he ok to eqt muffin.he kind of likes to
33CFKC	0	0	4	1	5.083333333	6.177777778	0.608081371	0.194504874	dosent like sports and likes sweets and healthy stuff to

Checking Data Leakage



Induced Features

- **Characteristic features (44) : 11 characteristics**

  personality inference, sweets, fruits and vegetables, healthy savory food, food, cosmetics, fashion, toys, gadgets and games, sports, music, arts and crafts

- **NLP features (16):**

  word count, sentence count, readability score, positive sentiment score, etc

Models Summary

Model Version	**Data Version**	**LLM Agent**	**Features**	**Train Accuracy**	**Test Accuracy**
V1	Data V2	Sonnet	Sub, profile, subject, SRS.RAW, FSR, BIS, avg_PE, free_response, LPA_Profile_grand_mean, LPA_Profile_ASD_only, NLP features	0.9159	0.8925
V2	Data V2	Sonnet	FSR, avg_PE, free_response, NLP features	0.9150	0.8699
V3	Data V2	Qwen	FSR, avg_PE, free_response, NLP features	0.9135	0.8699

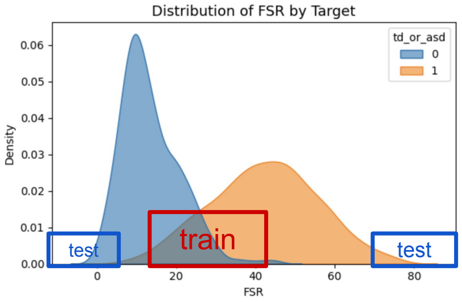
Models Summary

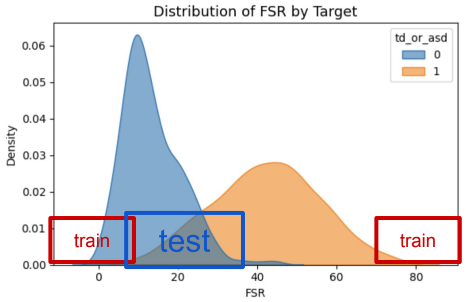
Model	Data File	LLM Agent	Features	Train Accuracy	Test Accuracy
V4	Data V2	Hugging Face Llama	FSR, avg_PE, free_response, NLP features	0.9072	0.8931
V5	Data V2, Data V1	Qwen	FSR, avg_PE, free_response, NLP features, slope features		
V6	Data V3	Qwen	FSR, avg_PE, free_response, NLP features, concept learning	0.8916	0.9062

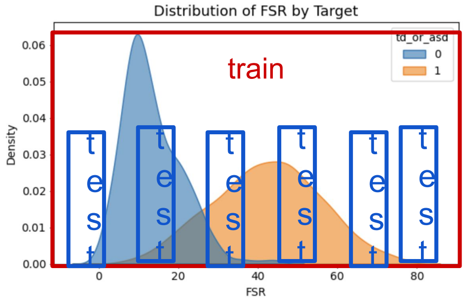
Models Summary

Model	Data File	LLM Agent	Features	Train Accuracy	Test Accuracy
V7-1	Data V3	Qwen	FSR, avg_PE, free_response, NLP features, concept learning	0.8954	0.8929
V7-2	Data V3	Qwen	Test on FSR overlap region	1.000	0.4727
V7-3	Data V3	Qwen	Train on FSR overlap region	0.8303	1.0000

FSR Overlap Test

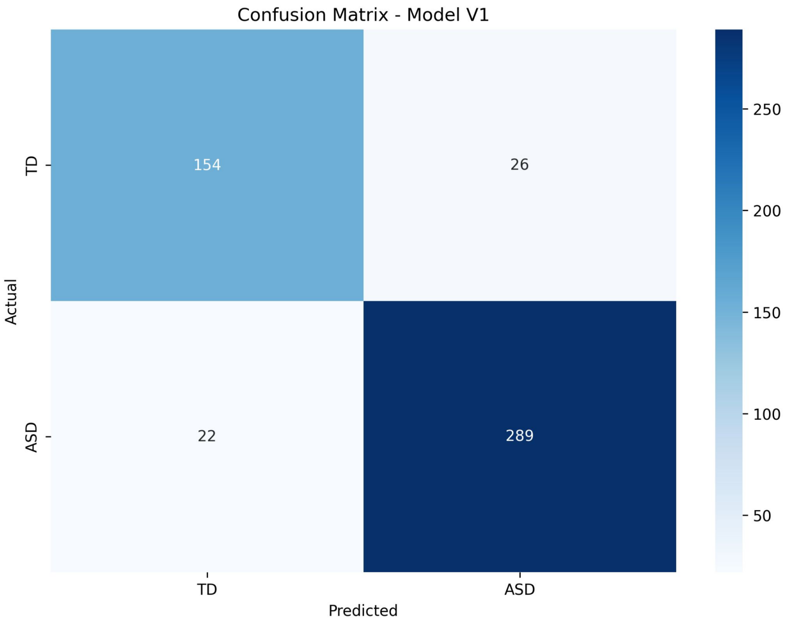






Results - V4 (Best Model) - on Data Aggregated

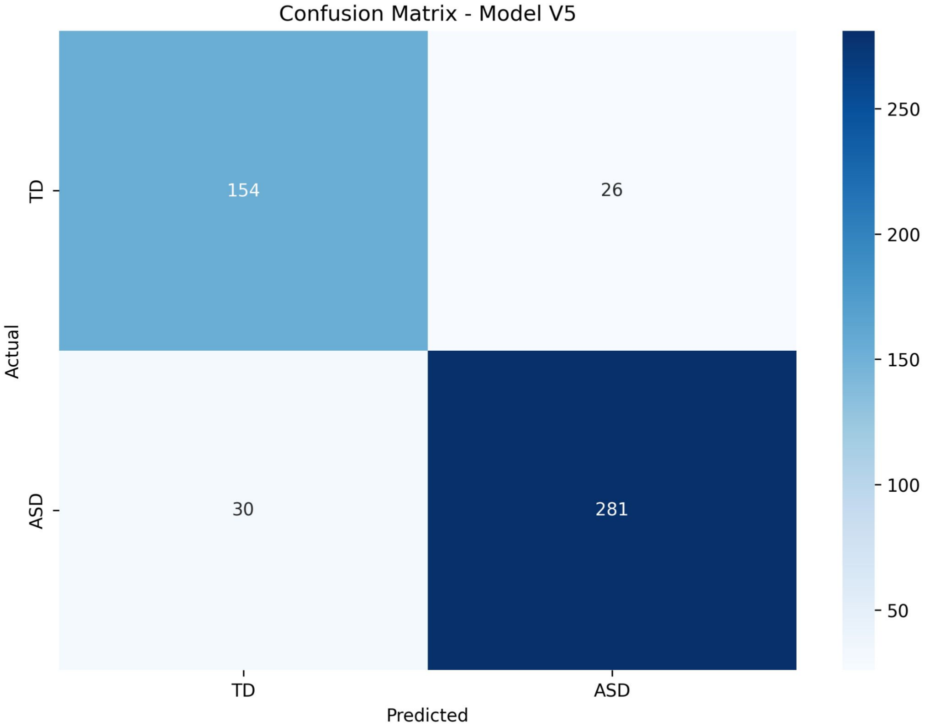
- Hugging Face Llama 3.2 for Characteristic Feature Extraction
-



Results - V5

- Using slope features calculated from “trial level data_v1_8.25.25.csv”

sub	mean_slope_ASDprof_norm	mean_slope_ASDprof_unif	mean_slope_TDprof_norm
235X4W	-0.019662301	0.002829671	-0.006179102
26NE6D	-0.007213784	-0.021714309	0.027223353
27YMQK	-0.068212834	NA	-0.010796393
2A7DFT	-0.00426993	0.095076331	0.02943333
2BMMAM	-0.057237514	NA	-0.045513839
2G3JNL	0.013183706	-0.04842936	-0.003017093
2GJdWc	0.030275468	-0.020900146	0.052073709
2HSJ5Z	0.008828889	0.035206622	0.008389702
2UAdZU	-0.051998461	0.059008475	0.011637296
2UK6c5	0.00524001	-0.032446855	-0.014401334
2WNcEH	0.011004821	-0.036493612	-0.01005928
33CFKC	-0.010973142	-0.001990427	-0.035362175
37EGLX	-0.025344699	0.024438646	-0.008510524
37S6ZU	0.015908832	-0.015749816	0.057525593

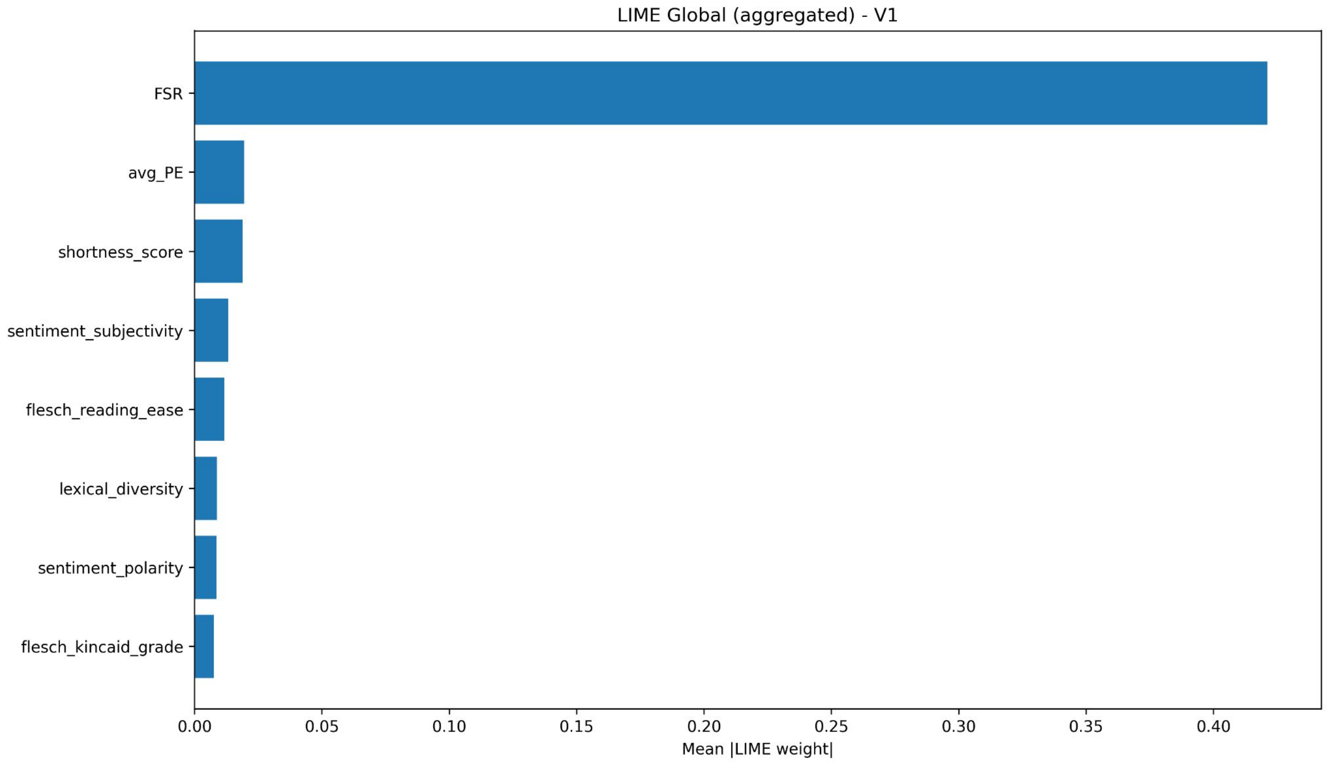


Results - V6

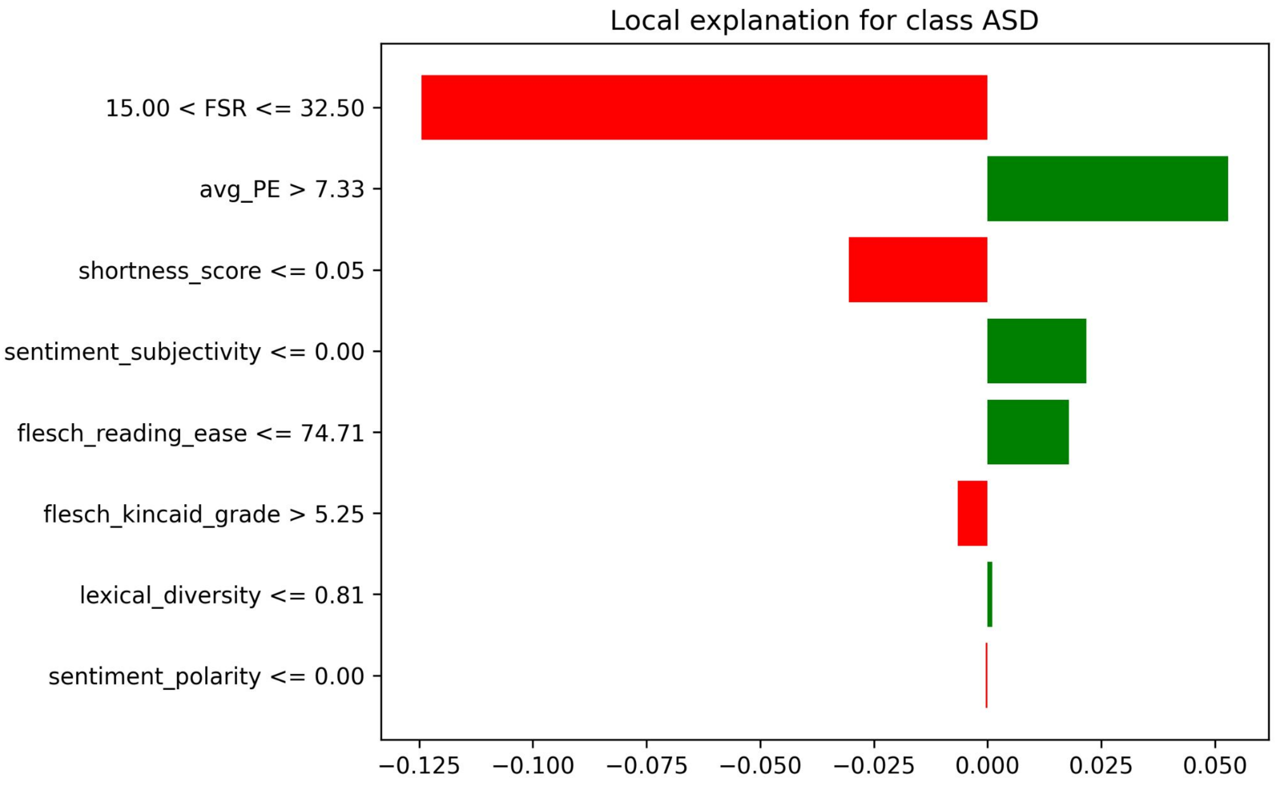
- Agent: Qwen/Qwen2.5-0.5B-Instruct
- Using concept learning features
- Optuna for xgboost_model

sub	td_or_asd	FSR	BIS	SRS.Raw	TDNorm_avg_PE	overall_avg_PE	TDnorm_concept_learning	overall_concept_learning	free_response_TDprof_norm
235X4W	0	5	4	10	5.716666667	7.111111111	-0.243313315	-0.004524569	He likes ice cream, kind of likes chess.
26NE6D	0	15	6	11	3.95	5.561111111	0.09796008	0.021768026	he doesn't like dice, skating, and exercising. really likes ch
27YMQK	0	NA	2.054054054	12	7.766666667	7.95	-0.021473482	0.020730875	he was normal
2A7DFT	0	7	0	7	4.266666667	4.938888889	-0.127769397	-0.151312665	She likes ice skating, fruit, the accordion, tasty food. She
2BMMAM	0	24	34	19	6.416666667	6.491666667	-0.339128115	-0.208756422	Quinn like's basketball she does not like game's she does
2G3JNL	0	11	9	20	4.633333333	5.822222222	0.060038511	0.012699918	Not a very pickey eater. Dislikes workouts. Detests most d
2GJdWc	0	3	12	18	4.566666667	5.644444444	0.138521821	0.069401618	Alex most notably liked more sophisticated foods, such as
2HSJ5Z	0	8	0	23	3.9	5.161111111	0.024022213	0.167810196	Alex likes food, especially fruits and sweets. He also really
2UAdZU	0	18	6	15	4.35	4.927777778	0.193282662	0.03618996	likes to eat pretty healthy
2UK6c5	0	17	18	20	5.183333333	5.755555556	0.246659019	0.069541374	He likes most musical instruments. He likes some ice ska
2WNcEH	0	41	48	58	5.5	6.388888889	-0.107478777	-0.14833702	he ok to plqy piano. he ok to eqt muffin.he kind of likes to
33CFKC	0	0	4	1	5.083333333	6.177777778	0.608081371	0.194504874	dosent like sports and likes sweets and heallthy stuff to

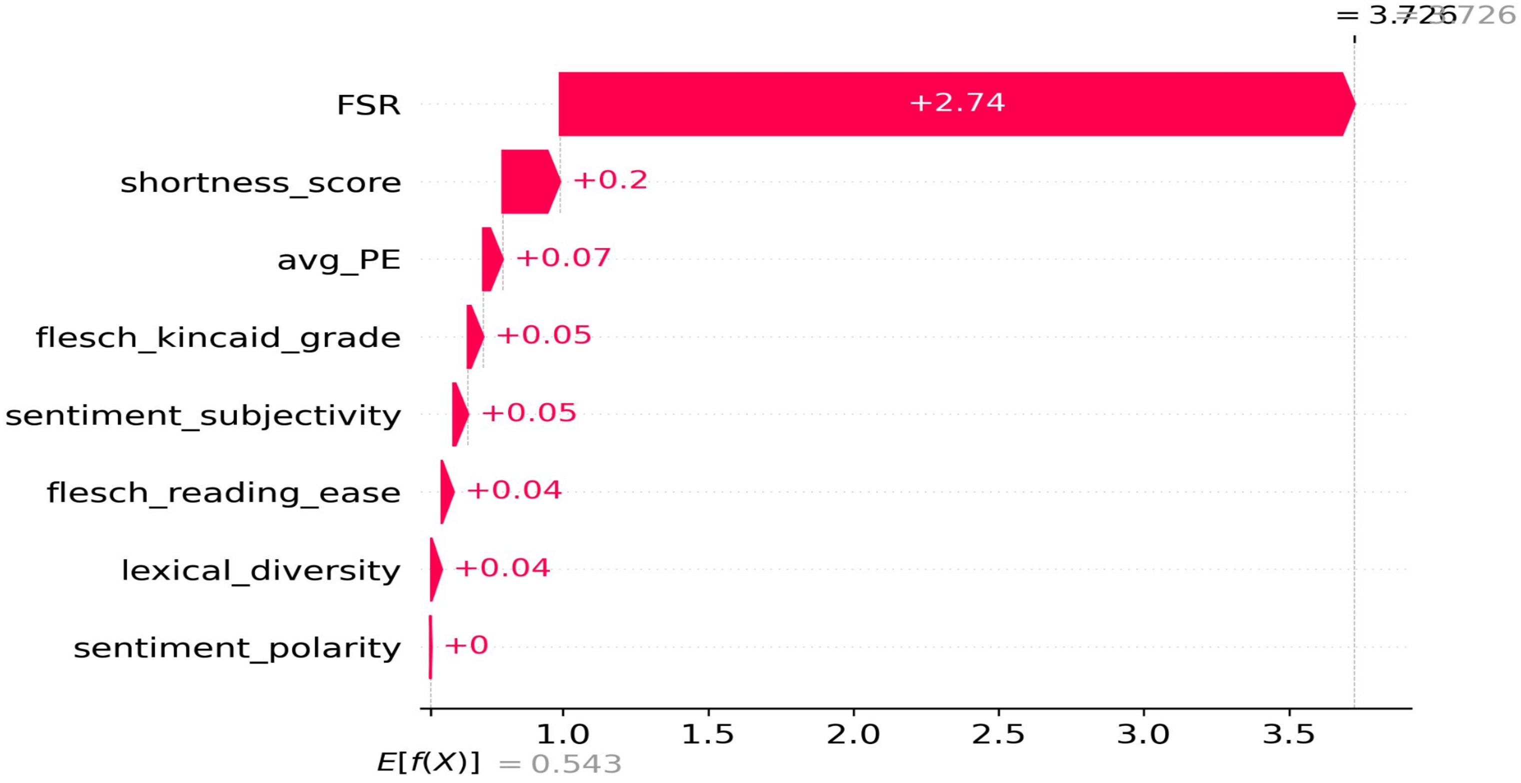
Global Explainability



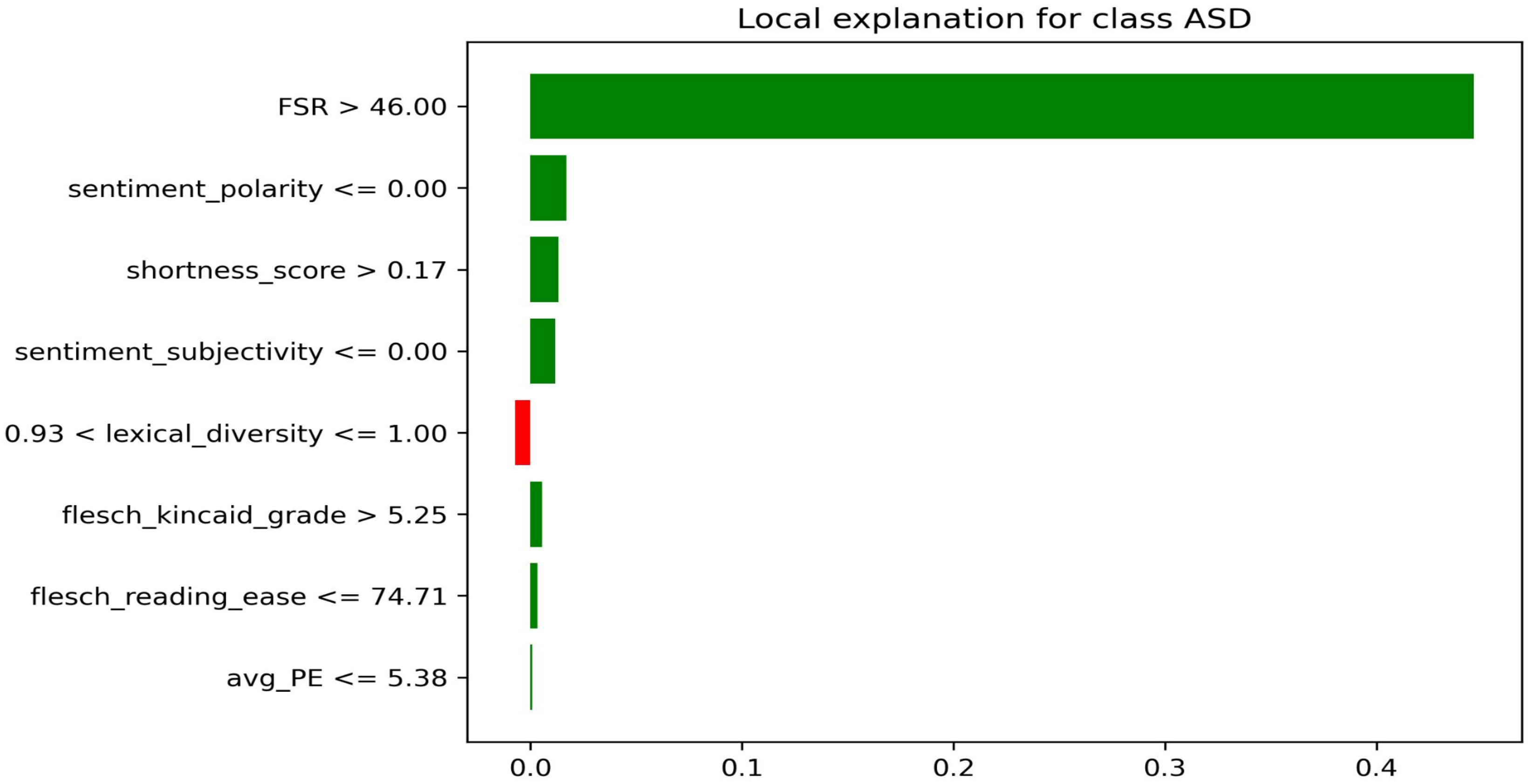
Local Explainability - row 1001 - Target 0 - TD



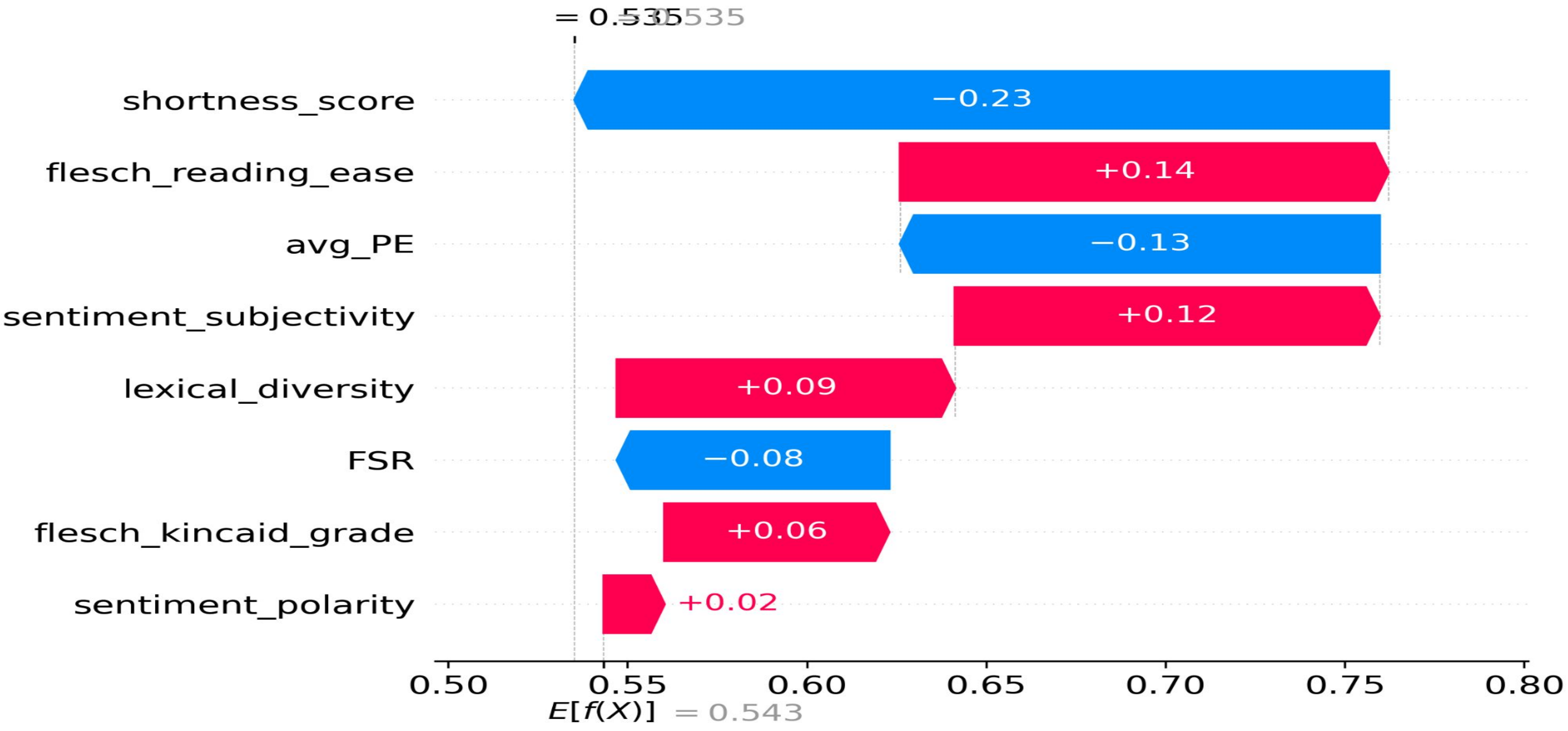
Local Explainability - row 1001 - Target 0 - TD



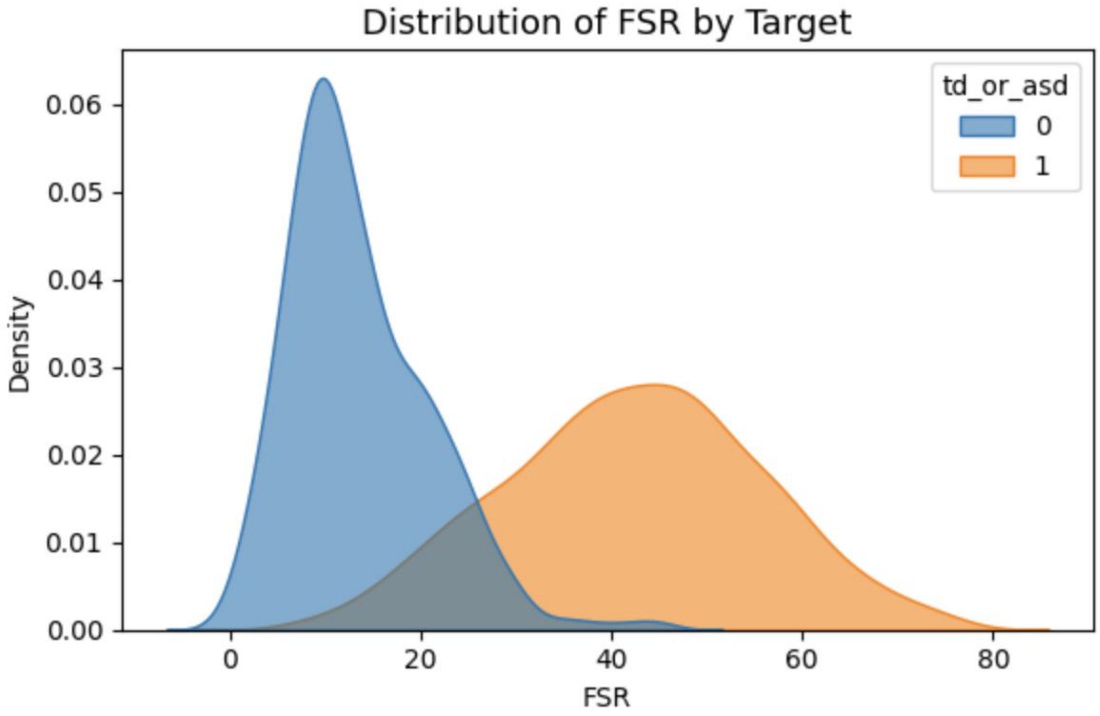
Local Explainability - row 1665 - Target 1 - ASD



Local Explainability - row 1665 - Target 1 - ASD



KDE plot of FSR by Target



FSR: Flexibility Scale Revised

- Higher scores indicate lower flexibility/ greater rigidity

Future Steps

1. Identify FSR overlap section between TD and ASD candidates and use that as Train/Test set to identify metrics
2. Identify Autism Subclusters - data approach
3. Modify Streamlit code to work with all versions and trials
4. Feature Importance Code needs to be modified for the TD ASD comparisons
5. Look at BERT Embeddings for text data to replace Characteristics