

Example 1

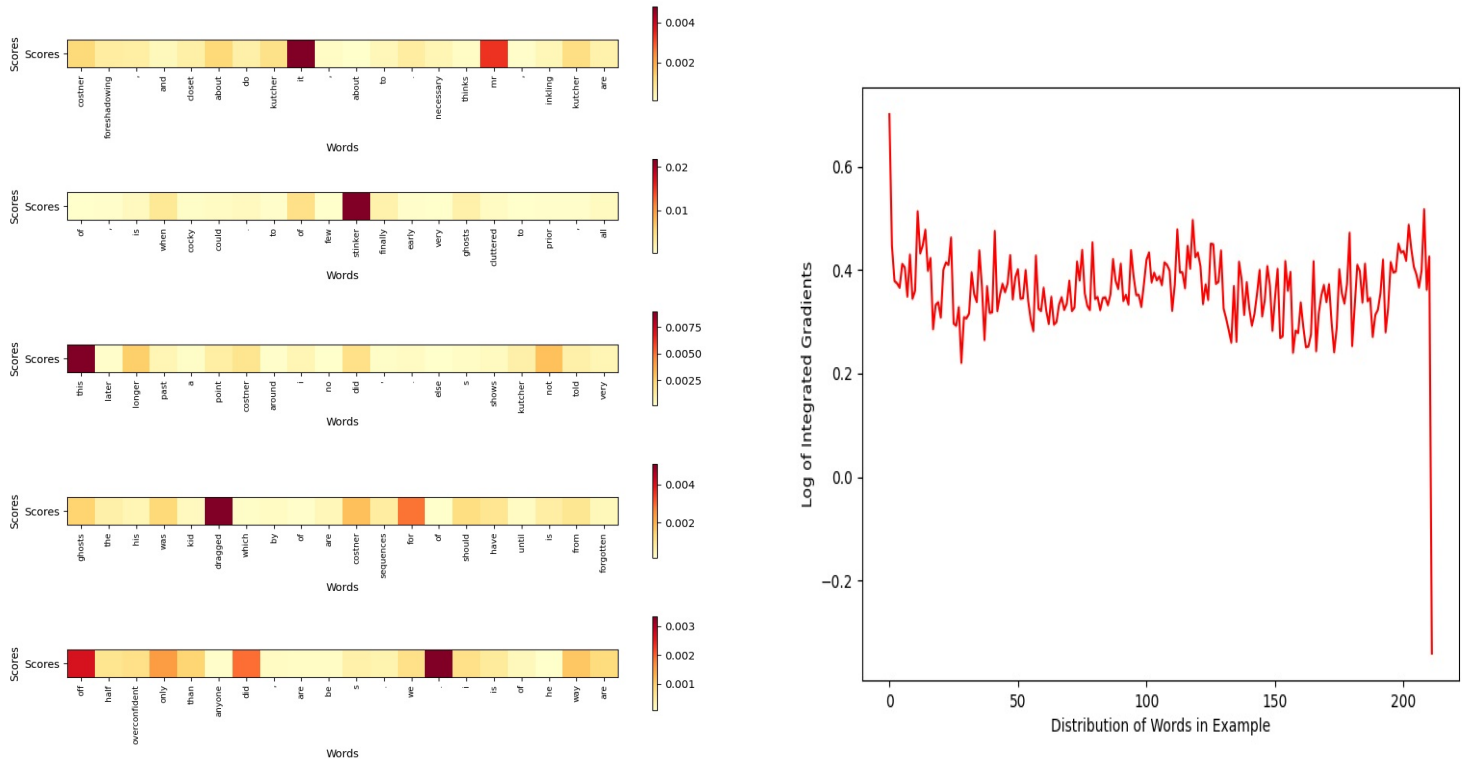
Example Text

Once again Mr. Costner has dragged out a movie for far longer than necessary. Aside from the terrific sea rescue sequences, of which there are very few I just did not care about any of the characters. Most of us have ghosts in the closet, and Costner's character are realized early on, and then forgotten until much later, by which time I did not care. The character we should really care about is a very cocky, overconfident Ashton Kutcher. The problem is he comes off as kid who thinks he's better than anyone else around him and shows no signs of a cluttered closet. His only obstacle appears to be winning over Costner. Finally when we are well past the half way point of this stinker, Costner tells us all about Kutcher's ghosts. We are told why Kutcher is driven to be the best with no prior inkling or foreshadowing. No magic here, it was all I could do to keep from turning it off an hour in.

Analysis Begins

Known Example Type:	Clean
Example original label:	0
Surrogate Model predicted label:	0
Confidence on Label:	0.99975914

Explainability Analysis



ATTENTION MAPS INTEGRATED GRADIENTS

Feature Importance Analysis of Adv Detector

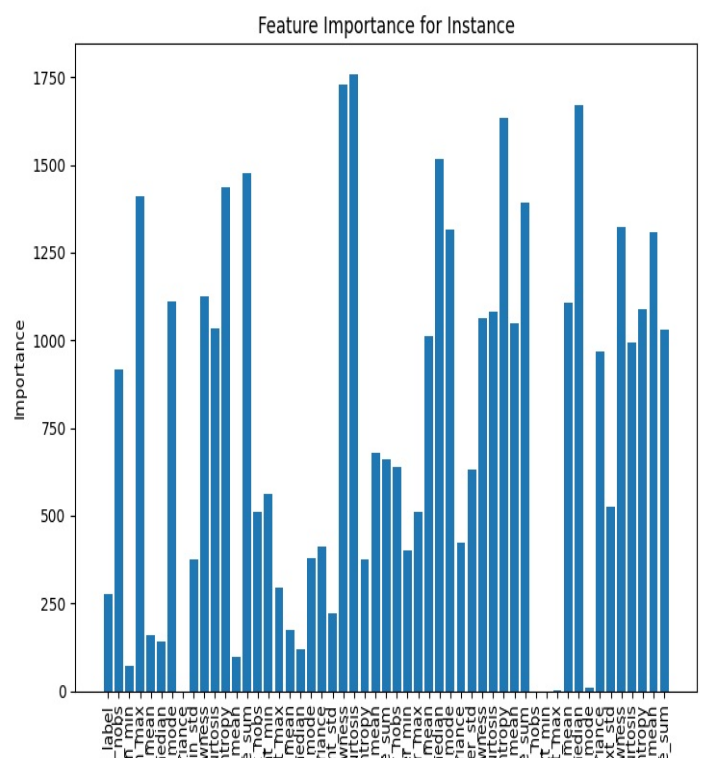
SHAP ANALYSIS LGBBOOST FEATURE IMPORTANCE

Detector Inference

Example is Detected as :Clean

Analysis finish

Example Text



SHAP ANALYSIS LGB00ST FEATURE IMPORTANCE

Detector Inference

Example is Detected as :Adversarial

Identification and Transformation

Identified Words:	['costner', 'vitale', 'wondrous', 'sea', 'sequential', 'zorg', 'peculiarity', 'poltergeists', 'costner', 'character', 'character', 'cocky', 'overconfident', 'ashton', 'kutcher', 'problem', 'drawer', 'appears', 'winning', 'costner', 'finally', 'halved', 'dimwit', 'costner', 'affirms', 'kutcher', 'told', 'kutcher', 'best', 'inkling', 'foreshadowing', 'aside', 'wondrous', 'sequential', 'poltergeists', 'closet', 'time', 'signs', 'obstacle', 'finally', 'past', 'halved', 'way', 'question', 'affirms', 'magic']
Possible Replacements:	{'wondrous': 'marvellous', 'sea': 'ocean', 'sequential': 'computation', 'peculiarity': 'shortcoming', 'poltergeists': 'commonplaces', 'character': 'protagonist', 'cocky': 'smug', 'overconfident': 'presumptuous', 'problem': 'serious', 'drawer': 'briefcase', 'appears': 'seems', 'winning': 'winner', 'finally': 'eventually', 'halved': 'trimmed', 'dimwit': 'plainspoken', 'affirms': 'tells'}
The predicted label after Transform:	0
Confidence on Label:	0.99636793
Human intervention is required:	NO
Message:	Converted to non-ADVERSARIAL EXAMPLE with newscore
Transformed Texted:	<p>whenever again mr . costner has dragged out a stills for far lange than vitale . aside from the marvellous ocean rescue computation , of which there are very few i just did not org about any of the shortcoming . most of us have commonplaces in the closet , and costner 's protagonist are realized early on , and then forgotten until much later , by which time i did not care . the character we should really care about is a very smug , presumptuous ashton kutcher . the serious is he comes off as kid who thinks he 's better than anyone else around him and shows no signs of a cluttered briefcase . his only obstacle seems to be winner over costner . eventually when we are well past the trimmed way question of this plainspoken , costner tells us all about kutcher 's ghosts . we are told why kutcher is driven to be the best with no prior inkling or foreshadowing . no magic here , it was all i could do to keep from turning it off an hour in .</p>

Example 2

Example Text

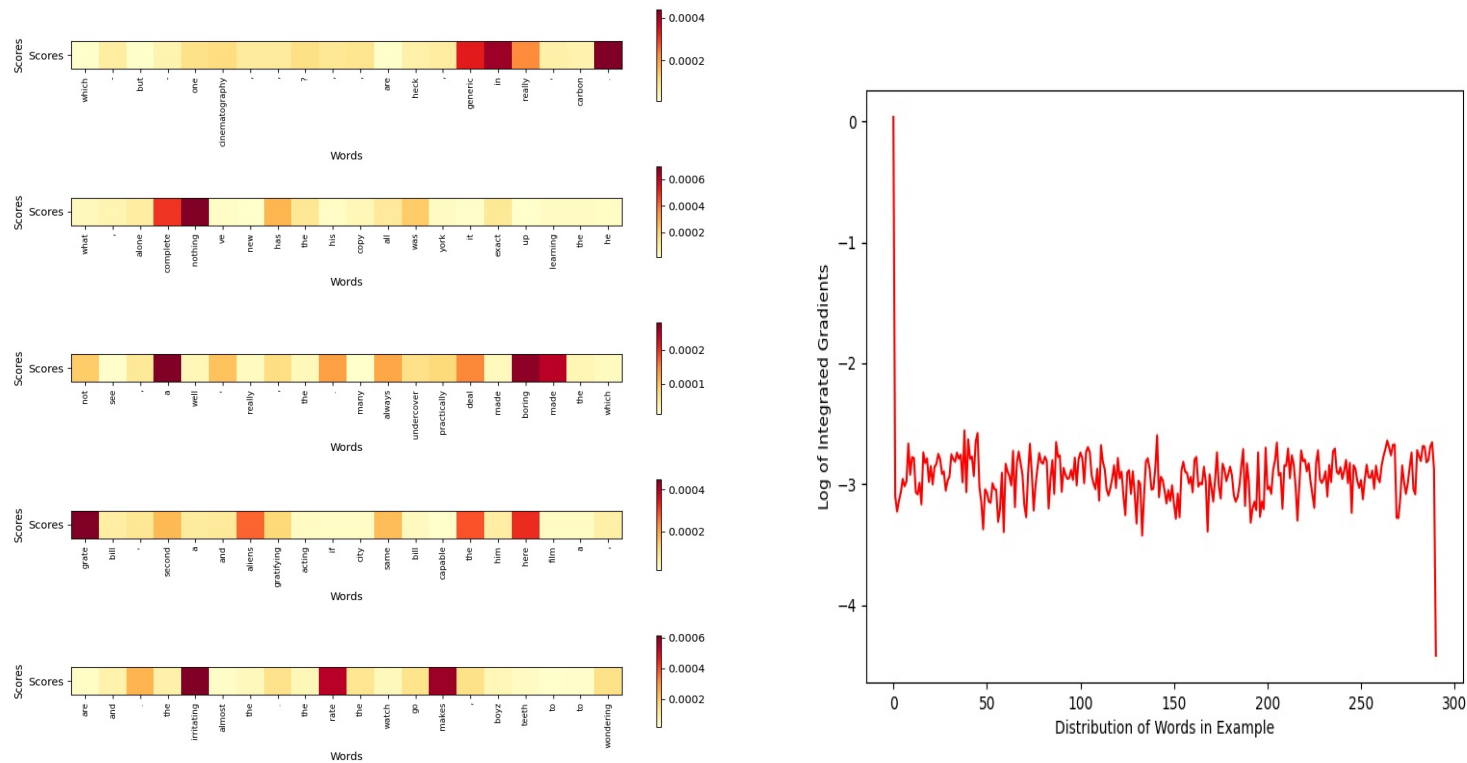
This is an example of why the majority of action films are the same. Generic and boring, there's really nothing worth watching here. A complete waste of the then barely-tapped talents of Ice-T and Ice Cube, who've each proven many times over that they are capable of acting, and acting well. Don't bother with this one, go see New Jack City, Ricochet or watch New York Undercover for Ice-T, or Boyz n the Hood, Higher Learning or Friday for Ice Cube and see the real deal. Ice-T's horribly cliched dialogue alone makes this film grate at the teeth, and I'm still wondering what the heck Bill Paxton was doing in this film? And why the heck does he always play the exact same character? From Aliens onward, every film I've seen with Bill Paxton has him playing the exact same irritating character, and at least in Aliens his character died, which made it somewhat gratifying...

Overall, this is second-rate action trash. There are countless better films to see, and if you really want to see this one, watch Judgement Night, which is practically a carbon copy but has better acting and a better script. The only thing that made this at all worth watching was a decent hand on the camera - the cinematography was almost refreshing, which comes close to making up for the horrible film itself - but not quite. 4/10.

Analysis Begins

Known Example Type:	Clean
Example original label:	0
Surrogate Model predicted label:	0
Confidence on Label:	0.99981695

Explainability Analysis



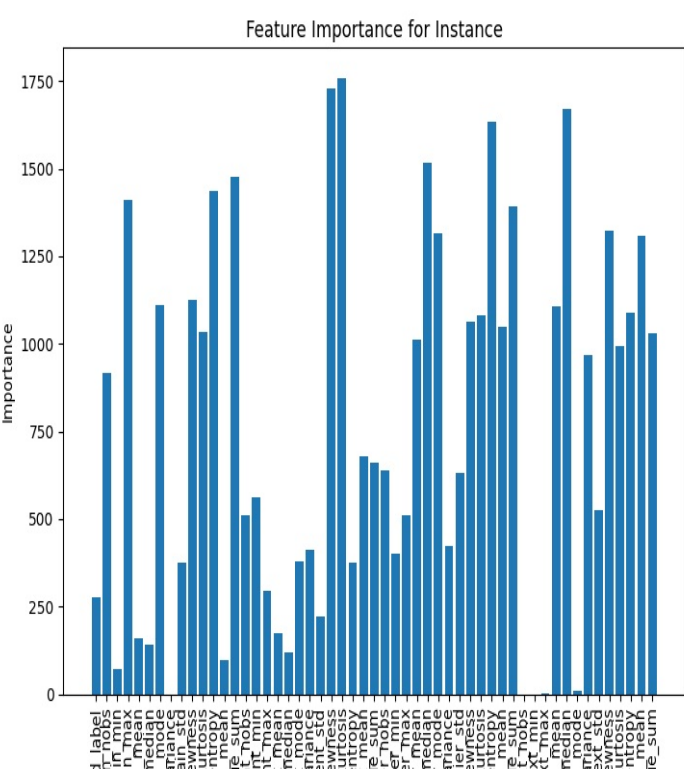
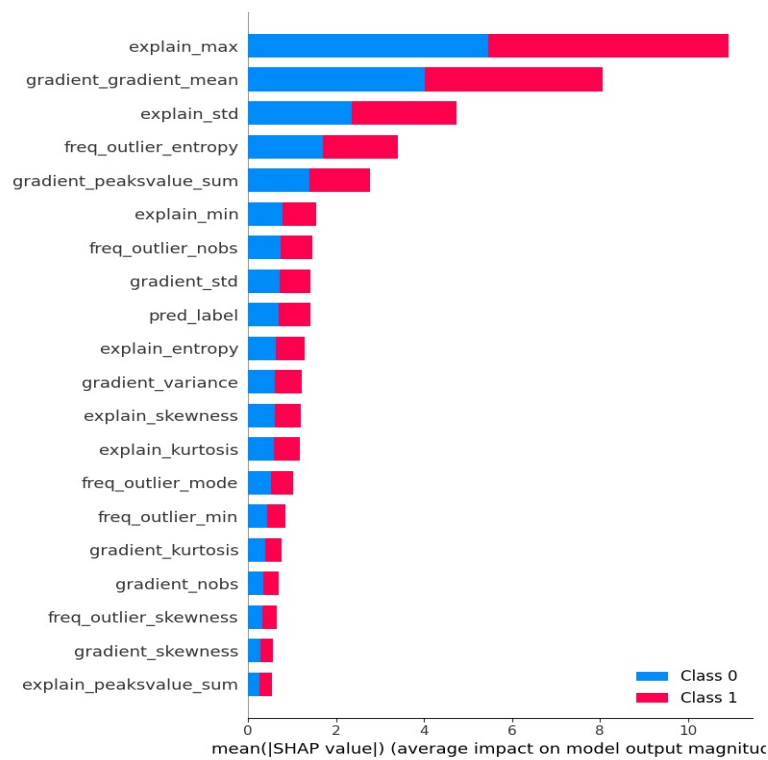
ATTENTION MAPS INTEGRATED GRADIENTS

Feature Importance Analysis of Adv Detector

SHAP ANALYSIS LGBBOOST FEATURE IMPORTANCE

Detector Inference

Example is Detected as :Clean



Analysis finish

Example Text

These is an example of why the majority of action films are the same. Generic and boring, there's really nothing worth watching here. A complete detritus of the then barely-tapped talents of Ice-T and Ice Cube, who've each proven many times over that they are capable of acting, and acting satisfactorily. Don't bother with this one, go see New Jack City, Ricochet or watch New York Undercover for Ice-T, or Boyz n the Hood, More Learning or Tonight for Icy Cube and admire the exact treat. Ice-T's terrifically cliched dialogue alone makes this cinematographers deflector at the teeth, and I'm still wondering what the scrawl Bill Paxton was put in this film? And why the heck does he always play the exact same character? Of Aliens onward, every film I've seen with Bill Paxton has him playing the exact same irritating character, and at least in Aliens his character died, which made it abundantly gratifying...

Overall, this is second-rate endeavor trash. There are countless better films to see, and if you really want to see this one, watch Judgement Night, which is practically a carbon copy but has better acting and a better script. The only gizmo that made this at all worth watching was a decent hand on the camera - the cinematography was almost refreshing, which comes close to making up for the horrible film itself - but not quite. 4/10.

Analysis Begins

Known Example Type:	Adversarial
Example original label:	0
Surrogate Model predicted label:	1
Confidence on Label:	0.5544128

Explainability Analysis

ATTENTION MAPS INTEGRATED GRADIENTS

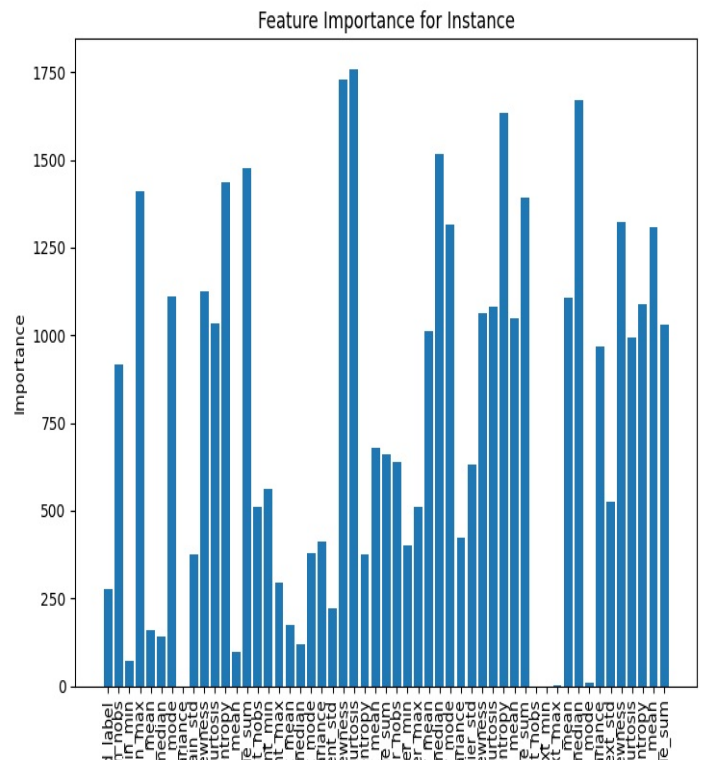
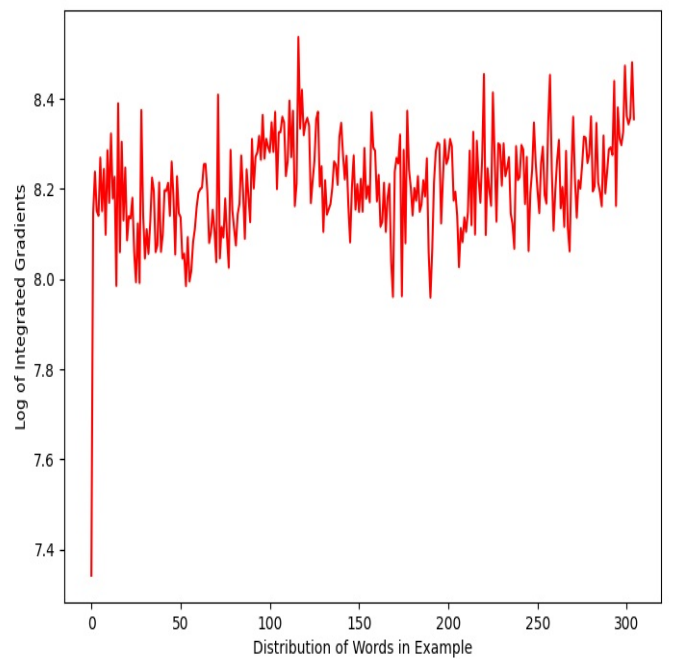
Feature Importance Analysis of Adv Detector

SHAP ANALYSIS LGBBOOST FEATURE IMPORTANCE

Detector Inference

Example is Detected as :Clean

Analysis finish



Example Text

The movie doesn't take more than five minutes to explain what is going on before we're already at the warehouse. There is not a single sympathetic character in this movie, except for the homeless guy, who is also the only one with half a brain.

The movie is filled with pointless violence and Walter Hills specialty: people falling through windows with glass flying everywhere. There is pretty much no plot and it is a big problem when you root for no-one. Everybody dies, except from Paxton and the homeless guy and everybody get what they deserve.

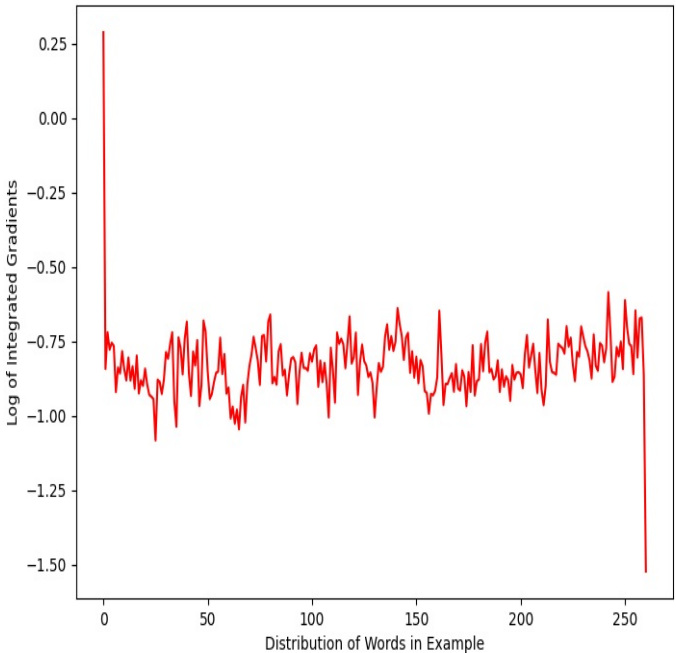
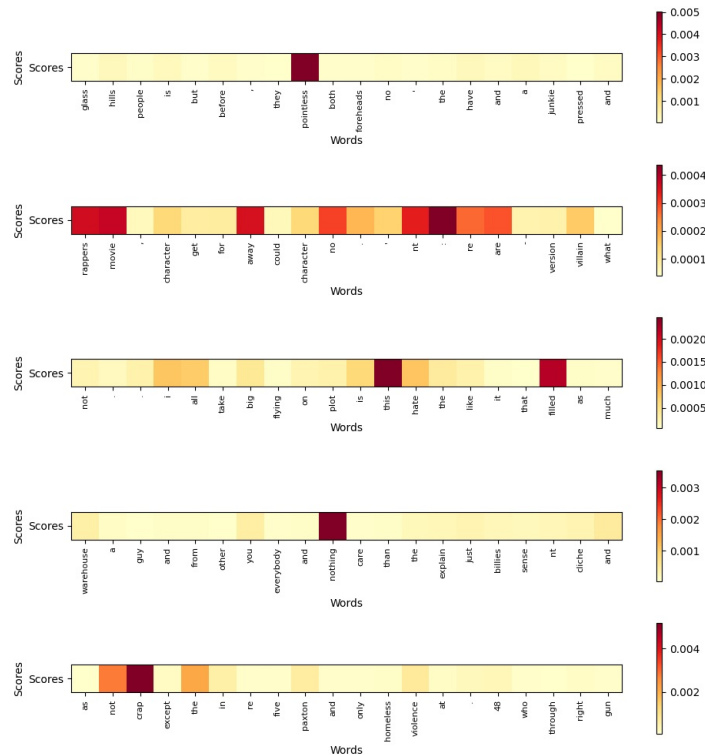
The only two black people that can act is the homeless guy and the junkie but they're actors by

Stay away from this crap and watch 48 hours 1 and 2 instead. At lest they have characters you care about, a sense of humor and nothing but real actors in the cast.

Analysis Begins

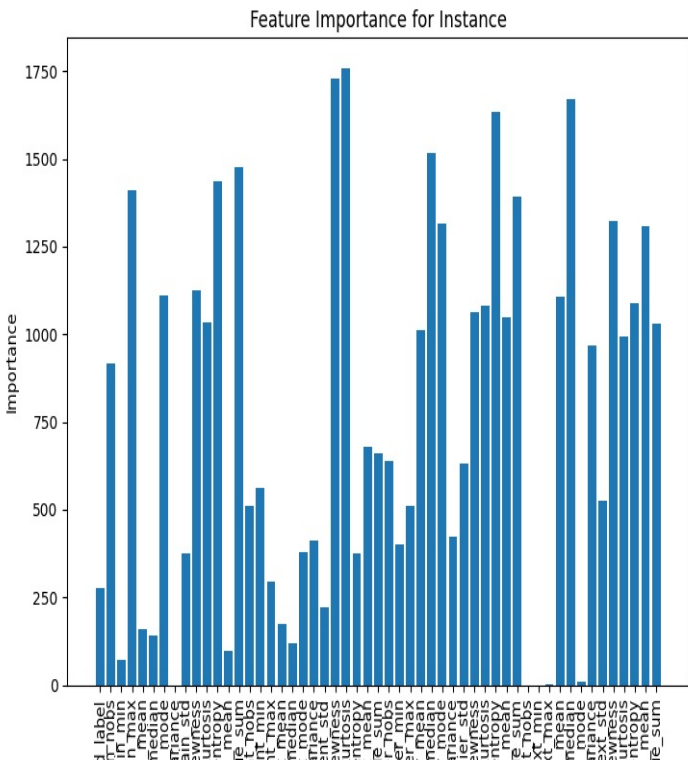
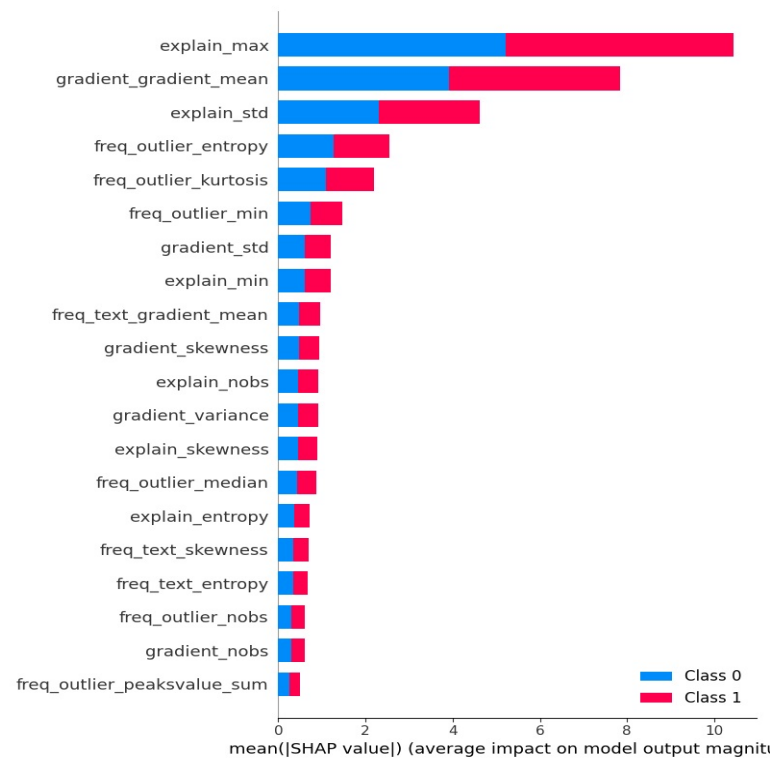
Known Example Type:	Clean
Example original label:	0
Surrogate Model predicted label:	0
Confidence on Label:	0.99971145

Explainability Analysis



ATTENTION MAPS INTEGRATED GRADIENTS

Feature Importance Analysis of Adv Detector



Detector Inference

Example is Detected as :Clean

Analysis finish

Example Text

First of all I hate those senseless rappers, who could'nt act if they had a gun pressed against their foreheads. All they do is curse and shoot each other and acting like clickÃ© version of traitors.

The movie doesn't take more than five minutes to explain what is going on before we're already at the warehouse There is not a single sympathetic character in this movie, except for the homeless guy, who is also the only one with half a brain.

Bill Paxton and William Sadler are both hill billies and Sadlers character is just as much a villain as the syndicates. I didn't like him right from the start.

The movie is filled with senseless violence and Walter Hills dedicated: people falling through windows with glass flying everywhere. There is pretty much no plot and it is a big problem when you root for no-one. Everybody dies, except from Paxton and the homeless guy and everybody get what they deserve.

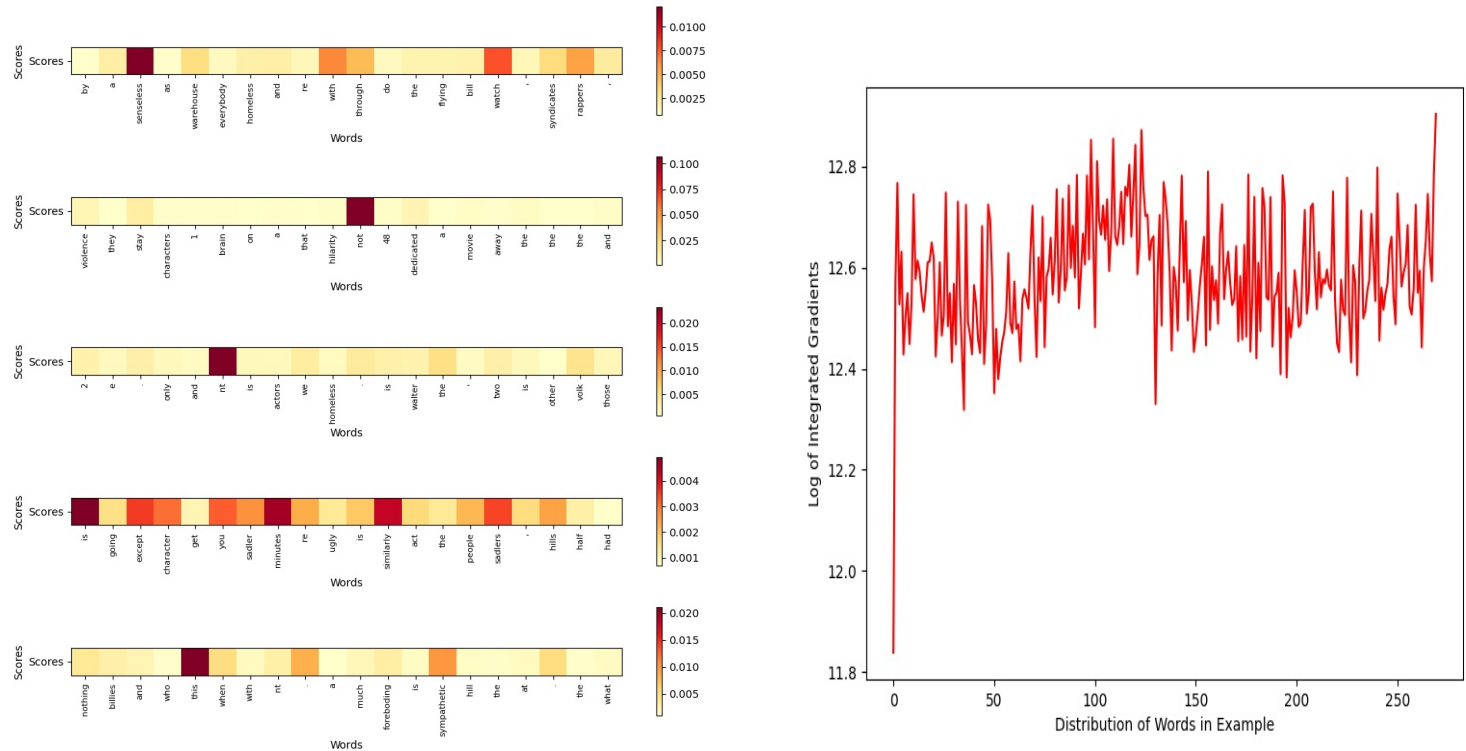
The only two black volk that can act is the homeless guy and the junkie but they're actors by profession, not annoying ugly brain dead rappers.

Stay away from this frak and watch 48 hours 1 and 2 similarly. Through mayest they have characters you care about, a foreboding of hilarity and nothing but truths officials in the castings.

Analysis Begins

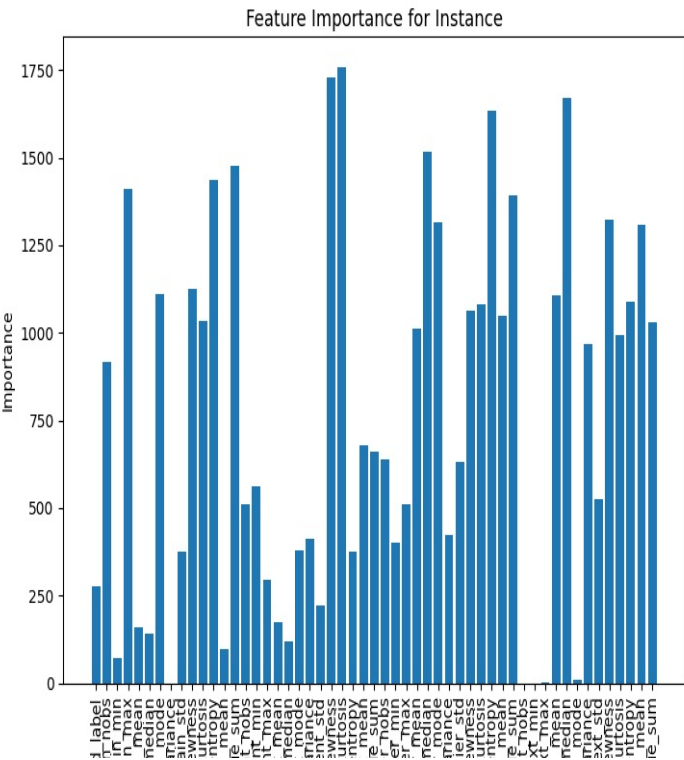
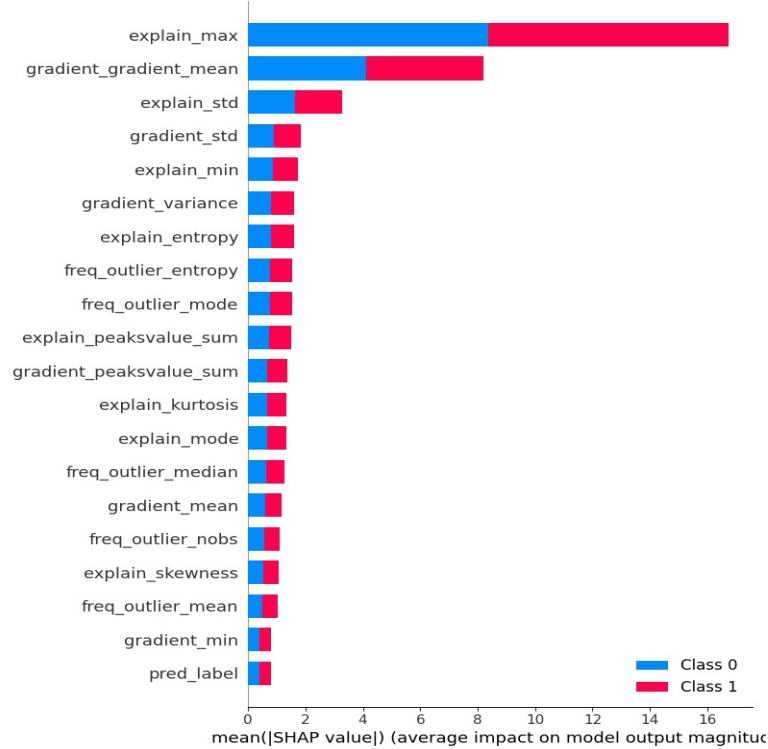
Known Example Type:	Adversarial
Example original label:	0
Surrogate Model predicted label:	1
Confidence on Label:	0.7168521

Explainability Analysis



ATTENTION MAPS INTEGRATED GRADIENTS

Feature Importance Analysis of Adv Detector



SHAP ANALYSIS LGBBOOST FEATURE IMPORTANCE

Detector Inference

Example is Detected as :Adversarial

Identification and Transformation

Identified Words:	['hate', 'senseless', 'rappers', 'foreheads', 'like', 'traitors', 'minutes', 'warehouse', 'single', 'sympathetic', 'sadler', 'sadlers', 'syndicates', 'like', 'filled', 'senseless', 'violence', 'dedicated', 'pretty', 'plot', 'volk', 'junkie', 'rappers', 'stay', 'frak', 'watch', 'mayest', 'foreboding', 'start', 'violence', 'hours', 'characters', 'care']
Possible Replacements:	{'hate': 'loathe', 'senseless': 'barbaric', 'rappers': 'rap', 'foreheads': 'chins', 'like': 'look', 'traitors': 'cowards', 'minutes': 'seconds', 'warehouse': 'store', 'single': 'tracks', 'sympathetic': 'supportive', 'syndicates': 'mafias', 'filled': 'room', 'violence': 'violent', 'pretty': 'really', 'plot': 'story', 'junkie': 'addict', 'stay': 'leave', 'frak': 'movie'}
The predicted label after Transform:	0
Confidence on Label:	0.9935493
Human intervention is required:	NO
Message:	Converted to non-ADVERSARIAL EXAMPLE with newscore
Transformed Texted:	first of all i loathe those barbaric rap , who couldn't act if they had a gun pressed against their chins . all they do is curse and shoot each other and acting look version of cowards . the movie does not take more than five seconds to explain what is going on before we 're already at the store there is not a tracks supportive character in this movie , except for the homeless guy , who is also the only one with half a brain . bill paxton and william sadler are both hill billies and sadlers character is just as much a villain as the mafias . i didn't like him right from the start . the movie is room with senseless violent and walter hills dedicated : people falling through windows with glass flying everywhere . there is really much no story and it is a big problem when you root for no - one . everybody dies , except from paxton and the homeless guy and everybody get what they deserve . the only two black vol that can act is the homeless guy and the addict but they 're actors by profession , not annoying ugly brain dead rappers . leave away from this movie and watch 48 hours 1 and 2 similarly . through mayest they have characters you care about , a foreboding of hilarity and nothing but truths officials in the castings .

