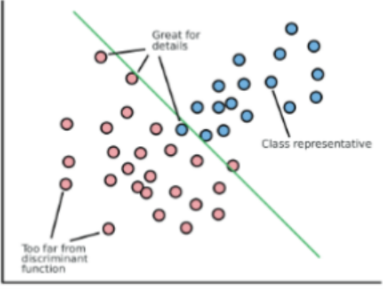

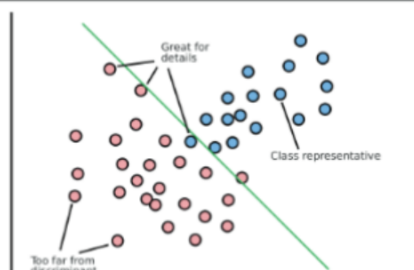


Multi-modal User Adaptation for Automatic Music Tagging

Yannis Vasilakis

2022

<ul style="list-style-type: none"> Research on music descriptors has been hypothesizing the existence of objective, universal musical properties But, each user might have a different definition for "Pop". What is yours? Can we measure/represent this "personal" opinion? Are users consistent with their definition? 	<p>This is not a Diva House song!</p> <p>Yup, it's more of a bubblegum dance house song..</p> <p>But they both sound like doom, tss, doom.</p> 
<h3>01 Background</h3> <p>Descriptors are inherently subjective and depend on:</p> <ul style="list-style-type: none"> Culture Education Experience <p>Machine Learning can be used to extract embedded spaces where every song is represented</p> <p>Each user will have a different distribution</p>	<p>Metalheads</p> <p>Scorpions Guns N' Roses</p> <p>Mortals</p> <p>Bruce Dickinson Anthrax Slay Scorpions Metallica Guns N' Roses Iron Maiden Edguy Blind Guardian Gamma-Ray Helloween</p> 
<h3>02 Technology</h3> <p>Training from scratch for each user is not computationally tractable</p> <ul style="list-style-type: none"> Users are not available for 40 hours Not every data point is informative <p>What if we could search and query only the informative data points?</p> <ul style="list-style-type: none"> Active Learning to the rescue! 	
<h3>03 Research Questions</h3> <ul style="list-style-type: none"> Can we evaluate the personalisation of definitions? <ul style="list-style-type: none"> Are users consistent with themselves? Can personally defined descriptors provide better: <ul style="list-style-type: none"> Accuracy (through embeddings) Profiling insights (user clustering) Are continuous defined descriptors useful? <ul style="list-style-type: none"> Can we measure "pop-ness"? Is this modality more useful? 	<h3>04 Future Work</h3> <ul style="list-style-type: none"> Hypothesis evaluation <ul style="list-style-type: none"> Questionnaire to evaluate users self-consistency with their descriptors Dataset formation <ul style="list-style-type: none"> Gather users, information about them, their personal descriptors Objective vs Subjective descriptors <ul style="list-style-type: none"> Find SOTA models for Genre Classification-Tags Compare their personalised accuracy

Finding

Nothing

Question

What is the least amount of song annotations needed to define a tag for a specific user? Do tags (or music descriptors in general) mean the same thing for different users?