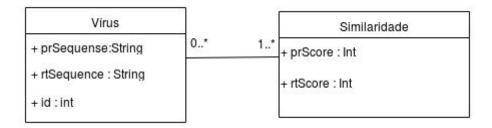
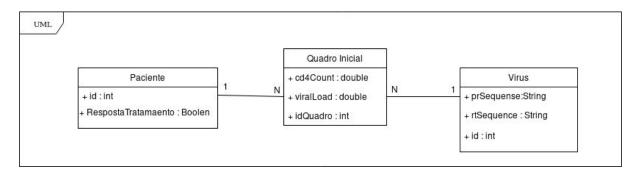
Predição de Efetividade do tratamento de HIV

Nesse projeto, pretendemos estudar a relação da estrutura viral e a efetividade do tratamento de HIV. Para isso, usaremos algoritmos de similaridade genética para estabelecer as similaridades entre os vírus (especificamente a Transcriptase Reversa (RT) e Protease (PR) que compõem seu material genético) e pretendemos utilizar Machine Learning para tentar prever a efetividade do tratamento e suas relações.

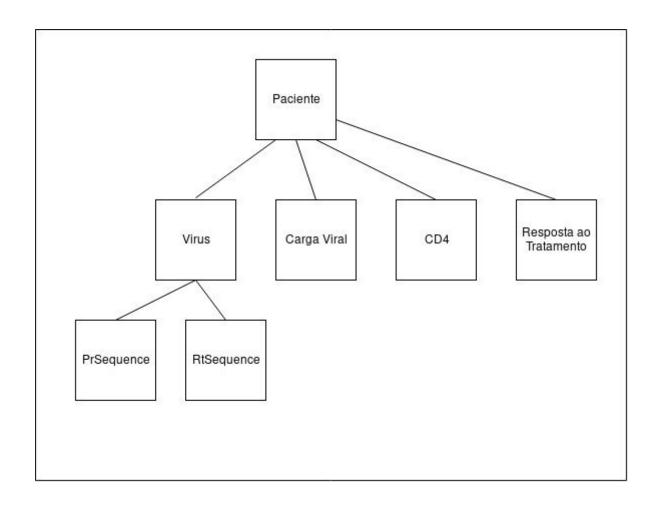
Nessa fase, foi necessário propor um modelo RDF e XML. Os modelos propostos foram úteis para conseguir fazer consultas em uma tabela única, poupando tempo de máquina em joins desnecessários. Para conseguir relacionar o problema proposto com um catálogo de possíveis relações do vírus, o qual num modelo relacional seria de difícil navegação entre as relações hierárquicas.

Modelo conceitual





Modelo Lógico XML



Consulta Xquery

Quantos pacientes responderam bem ao tratamento let \$root := doc('mydoc.xml') for \$p in (\$root) return count(\$p//root//Paciente[Resp=0])

ID dos pacientes ainda doentes let \$root := doc('mydoc.xml') for \$p in (\$root//Paciente) where \$p[Resp=0] return <Doente>{data(\$p/PatientID)}</Doente>

```
Pessoas que a contagem de cd4 menor que 50
let $root := doc('mydoc.xml')
for $p in ($root//Paciente)
where $p[CD4-t0 <50]
order by $p/PatientID
return <cd4>{data($p/PatientID)}</cd4>
Média de contagem cd4
let $root := doc('mydoc.xml')
for $p in ($root)
return avg($p//Paciente/CD4-t0)
Média de contagem VIRALLOAD
let $root := doc('mydoc.xml')
for $p in ($root)
return avg($p//Paciente/VL-t0)
                                               Consulta SPARQL
## Obter todos os tipos de infecções HIV
PREFIX rdf: <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>
PREFIX owl: <a href="http://www.w3.org/2002/07/owl#>">PREFIX owl: <a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#></a>
PREFIX meshv: <a href="http://id.nlm.nih.gov/mesh/vocab#">http://id.nlm.nih.gov/mesh/vocab#></a>
PREFIX mesh: <a href="http://id.nlm.nih.gov/mesh/">http://id.nlm.nih.gov/mesh/>
PREFIX mesh2015: <a href="http://id.nlm.nih.gov/mesh/2015/">http://id.nlm.nih.gov/mesh/2015/</a>
PREFIX mesh2016: <a href="http://id.nlm.nih.gov/mesh/2016/">http://id.nlm.nih.gov/mesh/2016/</a>
PREFIX mesh2017: <a href="http://id.nlm.nih.gov/mesh/2017/">http://id.nlm.nih.gov/mesh/2017/</a>
SELECT DISTINCT ?descriptor ?label
FROM <a href="http://id.nlm.nih.gov/mesh">http://id.nlm.nih.gov/mesh</a>
WHERE {
  mesh:D015658 meshv:treeNumber ?treeNum .
  ?childTreeNum meshv:parentTreeNumber+ ?treeNum .
  ?descriptor meshv:treeNumber ?childTreeNum .
  ?descriptor rdfs:label ?label .
}
ORDER BY ?label
LIMIT 10
```

```
## Obter todos os tipos de doenças sexualmente transmissíveis
PREFIX rdf: <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>>
PREFIX owl: <a href="http://www.w3.org/2002/07/owl#>"> http://www.w3.org/2002/07/owl#></a>
PREFIX meshv: <a href="http://id.nlm.nih.gov/mesh/vocab#">http://id.nlm.nih.gov/mesh/vocab#></a>
PREFIX mesh: <a href="http://id.nlm.nih.gov/mesh/">http://id.nlm.nih.gov/mesh/>
PREFIX mesh2015: <a href="http://id.nlm.nih.gov/mesh/2015/">http://id.nlm.nih.gov/mesh/2015/</a>
PREFIX mesh2016: <a href="http://id.nlm.nih.gov/mesh/2016/">http://id.nlm.nih.gov/mesh/2016/</a>
PREFIX mesh2017: <a href="http://id.nlm.nih.gov/mesh/2017/">http://id.nlm.nih.gov/mesh/2017/</a>
SELECT DISTINCT ?descriptor ?label
FROM <a href="http://id.nlm.nih.gov/mesh">http://id.nlm.nih.gov/mesh</a>>
WHERE {
           ?descriptor rdfs:label ?label .
           ?descriptor meshv:broaderDescriptor mesh:D015229
ORDER BY ?label
## Obter todas as annotations das infecções HIV
PREFIX rdf: <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>>
PREFIX owl: <a href="http://www.w3.org/2002/07/owl#>"> http://www.w3.org/2002/07/owl#></a>
PREFIX meshv: <a href="http://id.nlm.nih.gov/mesh/vocab#">http://id.nlm.nih.gov/mesh/vocab#></a>
PREFIX mesh: <a href="http://id.nlm.nih.gov/mesh/">http://id.nlm.nih.gov/mesh/>
PREFIX mesh2015: <a href="http://id.nlm.nih.gov/mesh/2015/">http://id.nlm.nih.gov/mesh/2015/</a>
PREFIX mesh2016: <a href="http://id.nlm.nih.gov/mesh/2016/">http://id.nlm.nih.gov/mesh/2016/</a>
PREFIX mesh2017: <a href="http://id.nlm.nih.gov/mesh/2017/">http://id.nlm.nih.gov/mesh/2017/</a>
SELECT DISTINCT ?label ?annotation
FROM <a href="http://id.nlm.nih.gov/mesh">http://id.nlm.nih.gov/mesh</a>
WHERE {
  mesh:D015658 meshv:treeNumber ?treeNum .
  ?childTreeNum meshv:parentTreeNumber+ ?treeNum .
  ?descriptor meshv:treeNumber ?childTreeNum .
  ?descriptor rdfs:label ?label .
  ?descriptor meshv:annotation ?annotation
}
ORDER BY ?label
LIMIT 10
```

```
## Obter todas as doenças relacionadas a slow vírus
PREFIX rdf: <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>>
PREFIX owl: <a href="http://www.w3.org/2002/07/owl#>">PREFIX owl: <a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#></a>
PREFIX meshv: <a href="http://id.nlm.nih.gov/mesh/vocab#">http://id.nlm.nih.gov/mesh/vocab#></a>
PREFIX mesh: <a href="http://id.nlm.nih.gov/mesh/">http://id.nlm.nih.gov/mesh/>
PREFIX mesh2015: <a href="http://id.nlm.nih.gov/mesh/2015/">http://id.nlm.nih.gov/mesh/2015/</a>
PREFIX mesh2016: <a href="http://id.nlm.nih.gov/mesh/2016/">http://id.nlm.nih.gov/mesh/2016/</a>
PREFIX mesh2017: <a href="http://id.nlm.nih.gov/mesh/2017/">http://id.nlm.nih.gov/mesh/2017/</a>
 SELECT DISTINCT ?descriptor ?label
 FROM <a href="http://id.nlm.nih.gov/mesh">http://id.nlm.nih.gov/mesh</a>
 WHERE {
     mesh:D012897 meshv:treeNumber ?treeNum .
     ?childTreeNum meshv:parentTreeNumber+ ?treeNum .
     ?descriptor meshv:treeNumber ?childTreeNum .
     ?descriptor rdfs:label ?label .
 }
 ORDER BY ?label
LIMIT 10
## Obter todas as síndromes de deficiência imunológica
PREFIX rdf: <a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
PREFIX rdfs: <a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>
PREFIX owl: <a href="http://www.w3.org/2002/07/owl#>"> PREFIX owl: <a href="http://www.w3.org/2002/07/owl#"> PREFIX owl: <a href="http://www.wa.org/2002/07/owl#"> PREFIX owl: <a href="http://www.wa.org/2002/07/
PREFIX meshv: <a href="http://id.nlm.nih.gov/mesh/vocab#">http://id.nlm.nih.gov/mesh/vocab#></a>
PREFIX mesh: <a href="http://id.nlm.nih.gov/mesh/">http://id.nlm.nih.gov/mesh/>
PREFIX mesh2015: <a href="http://id.nlm.nih.gov/mesh/2015/">http://id.nlm.nih.gov/mesh/2015/</a>>
PREFIX mesh2016: <a href="http://id.nlm.nih.gov/mesh/2016/">http://id.nlm.nih.gov/mesh/2016/</a>
PREFIX mesh2017: <a href="http://id.nlm.nih.gov/mesh/2017/">http://id.nlm.nih.gov/mesh/2017/</a>
 SELECT DISTINCT ?descriptor ?label
 FROM <a href="http://id.nlm.nih.gov/mesh">http://id.nlm.nih.gov/mesh</a>
 WHERE {
     mesh:D007153 meshv:treeNumber ?treeNum .
     ?childTreeNum meshv:parentTreeNumber+ ?treeNum .
     ?descriptor meshv:treeNumber ?childTreeNum .
     ?descriptor rdfs:label ?label .
 }
 ORDER BY ?label
LIMIT 10
```

```
Exemplo de tabela XML convertida no <a href="http://convertcsv.com/csv-to-xml.htm">http://convertcsv.com/csv-to-xml.htm</a>
<?xml version="1.0" encoding="UTF-8"?>
<root>
 <Paciente>
```

<PatientID>1</PatientID> <Resp>0</Resp>

<PRSeq>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCCCAATAAGGATAGGGGGGGCAACTAAAGGAAGC YCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGACATGGAGTTGCCAGGAAGATGGAAACCAAAAAT GATAGGGGGAATTGGAGGTTTTATCAAAGTAARACAGTATGATCAGRTACCCATAGAAATCTATGGACATAAA GCTGTAGGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGCTTGGTT GCACTTTAAATTTY</PRSeq>

CAATGGCCATTGACAGAAGAAAAAAAAAAAGCATTAGTAGAAATTTGYACAGAAATGGAAAAGGAAGGGAAA ATTTCAAAAATTGGGCCTGAAAATCCATATAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTACATG GAGAAAATTAGTAGATTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAAYTAGGAATACCA CATCCCGCWGGGTTAAAAAAGAAYAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTCTCAGTTCCM TTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGAT ATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCT TTTRGAAATAGAACAGCATAGAACAAAAATAGAGGAACTGAGACAACATCTGTCAAGGTGGGGGTTTACCAC ACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGCTATGAACTCCATCCTGATAAATGGAC AGTACAGCCTATAGTTCTGCCAGAAAAAGATAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGGAAGTT ACTAACAGAAATAATACCACTAACAAGAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAAAATTCTAAAAGA ACCAGTACATGGAGTGTATTATGATCCAACAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGC</R TSeq>

```
<VL-t0>4.3</VL-t0>
       <CD4-t0>145</CD4-t0>
</Paciente>
<Paciente>
       <PatientID>2</PatientID>
       <Resp>0</Resp>
```

<PRSeq>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCGCAATAAAGATAGGGGGGGCAACTAAAGGAAGC TCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGACATGGAATTGCCAGGAAGATGGAAACCAAAAAT AATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACCCATAGAAATCTGTGGACATAAA GTTATAAGTACAGTATTAATAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGATGACTCAGCTTGGTT GCACTTTAAATTTT</PRSeq>

CAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGARATGGAAARGGARGGGAAA ATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCGGTATTTGTCATAAAGAAAAAGGACAGTACTAAGT GGAGAAAGTAGTAGATTTCAGAGAACTTAATAAAAGAACTCAAGACTTCTGGGAAGTTCAATTAGGGATACC ACATCCCGCAGGGWTAAAAAAGAATAAATCAGTAACAGTATTGGATGTGGGTGATGCATACTTTTCAGTTCC ATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATT

```
<VL-t0>3.6</VL-t0>
<CD4-t0>224</CD4-t0>
</Paciente>

<Paciente>
<PatientID>3</PatientID>
<Resp>0</Resp>
```

<PRSeq>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCGCAATAAAGGTAGGGGGGGCAACTAAAAGAAGC TCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGACATGAGTTTGCCAGGAAAATGGAAACCAAAAAT GATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCGTAGAAATCTGTGGACATAA AGCTATAGGTACAGTATTRGTAGGACCTACACCTGTYAACATAATTGGAAGAAATCTGTTRACTCAGATTGGY TGCACTTTAAATTTT</PRSeq>

```
<VL-t0>3.2</VL-t0>
<CD4-t0>1017</CD4-t0>
</Paciente>

<Paciente>

<PatientID>4</PatientID>
<Resp>0</Resp>
```

<PRSeq>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCGCAATAAGGATAGGGGGGGCAACTAAAGGAAGC TCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGACATGAATTTGCCAGGGAAATGGAAACCAAAAAT GATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGAGGAGATACCCATAGAGATCTGTGGRCATAA AGCTATAGGTACAGTATTAGTAGGACCAACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGC TGCACTTTAAATTTT</PRSeq>

```
<VL-t0>5.7</VL-t0>
<CD4-t0>206</CD4-t0>
</Paciente>

<Paciente>
<PatientID>5</PatientID>
<Resp>0</Resp>
```

<PRSeq>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCGCAGTAAAGATAGGGGGGGCAACTAAAGGAAGC TCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGACTTTGCCAGGAAGATGGAAACCAAAAAT GATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGGTACCCATAGAAATCTGTGGACATAA AGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGT TGCACTTTAAATTTC</PRSeq>

```
<VL-t0>3.5</VL-t0>
<CD4-t0>572</CD4-t0>
</Paciente>

<Paciente>

<PatientID>6</PatientID>
<Resp>0</Resp>
```

<PRSeq>CCTCAAATCACTCTTTGGCAACGACCCCTCGTCTCAATAAAGATAGGGGGGGCAACTAAAAGAAGC CCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAGATGGAGACCAAAAAT GATAGGGGGAATTGGAGGTTTTATCAAAGTAAAACAGTATGATCAGGTACCCATAGAAATCTGTGGACATAA AACTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGCTGACTCAGCTTGGT TGCACTTTAAATTTT</PRSeq> <VL-t0>3.9</VL-t0> <CD4-t0>239</CD4-t0> </Paciente>

</root>