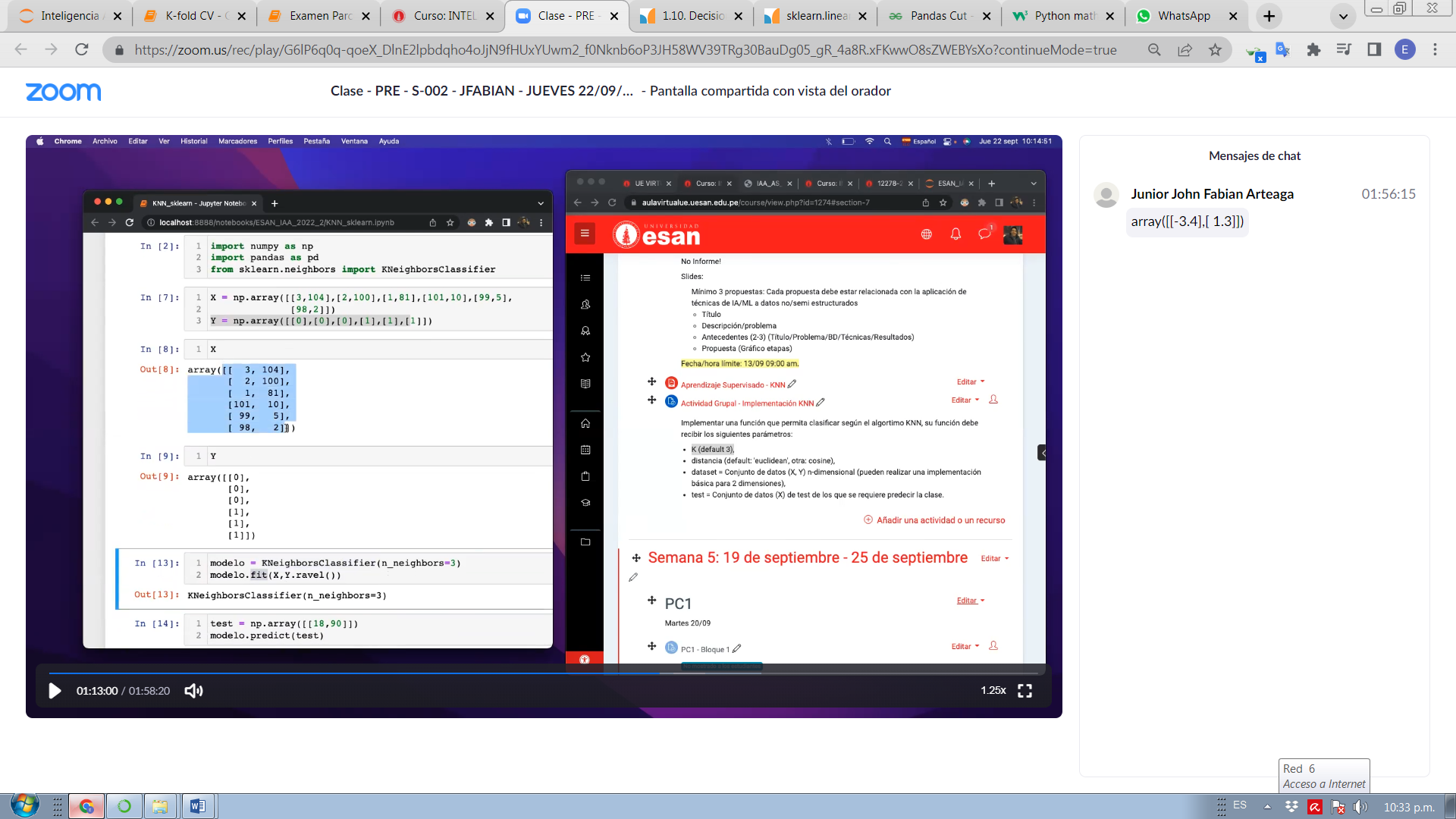
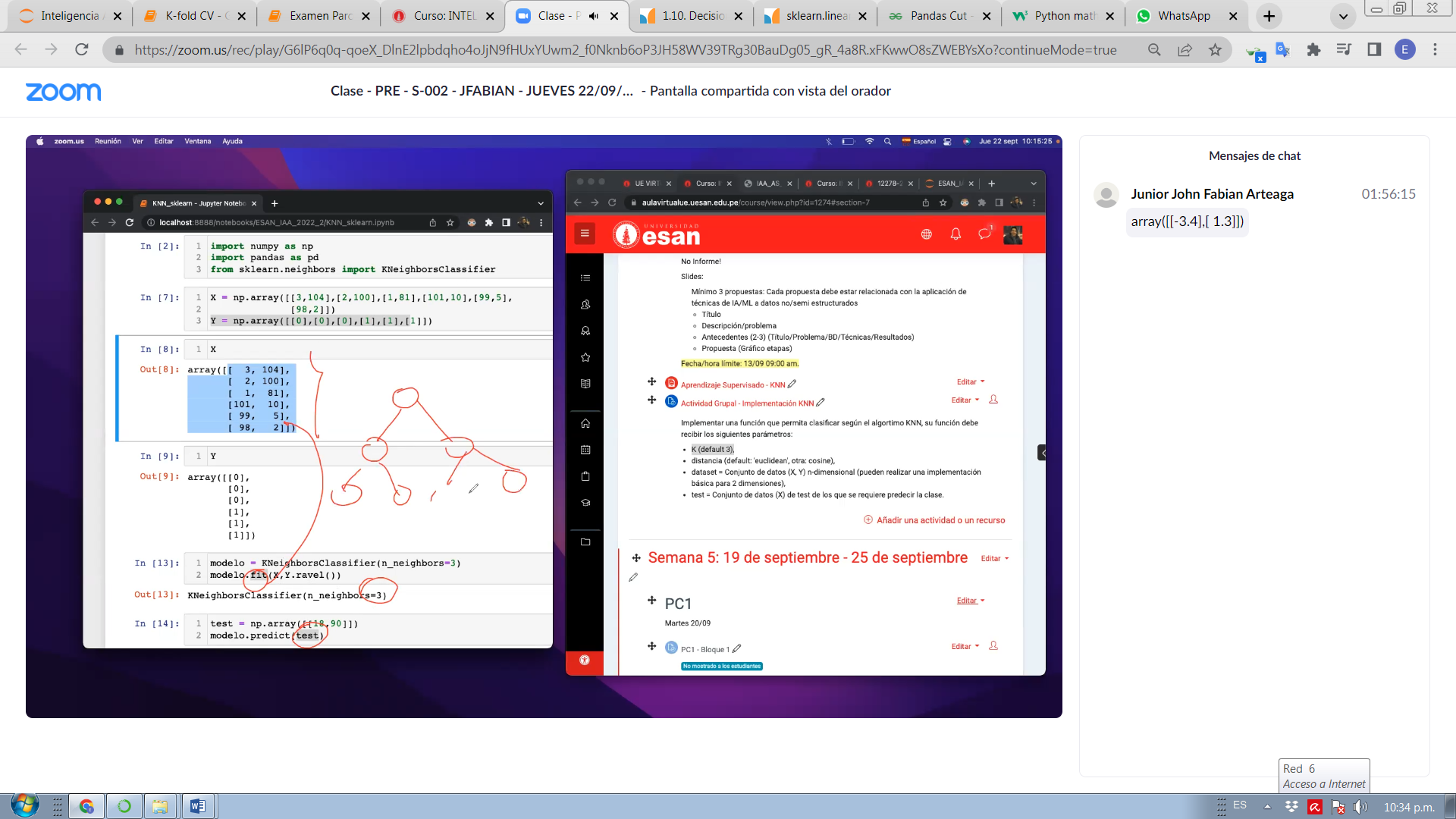
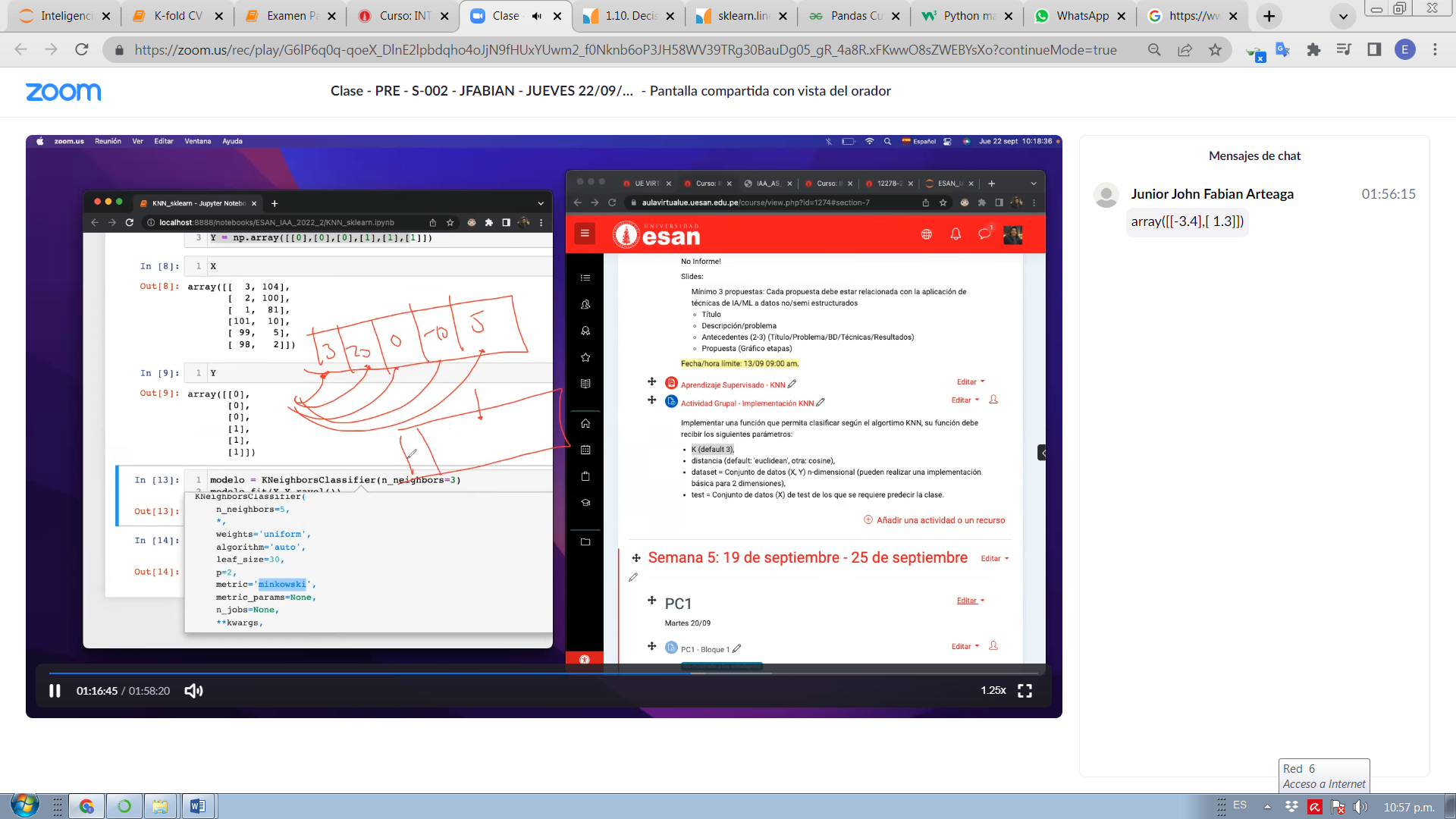
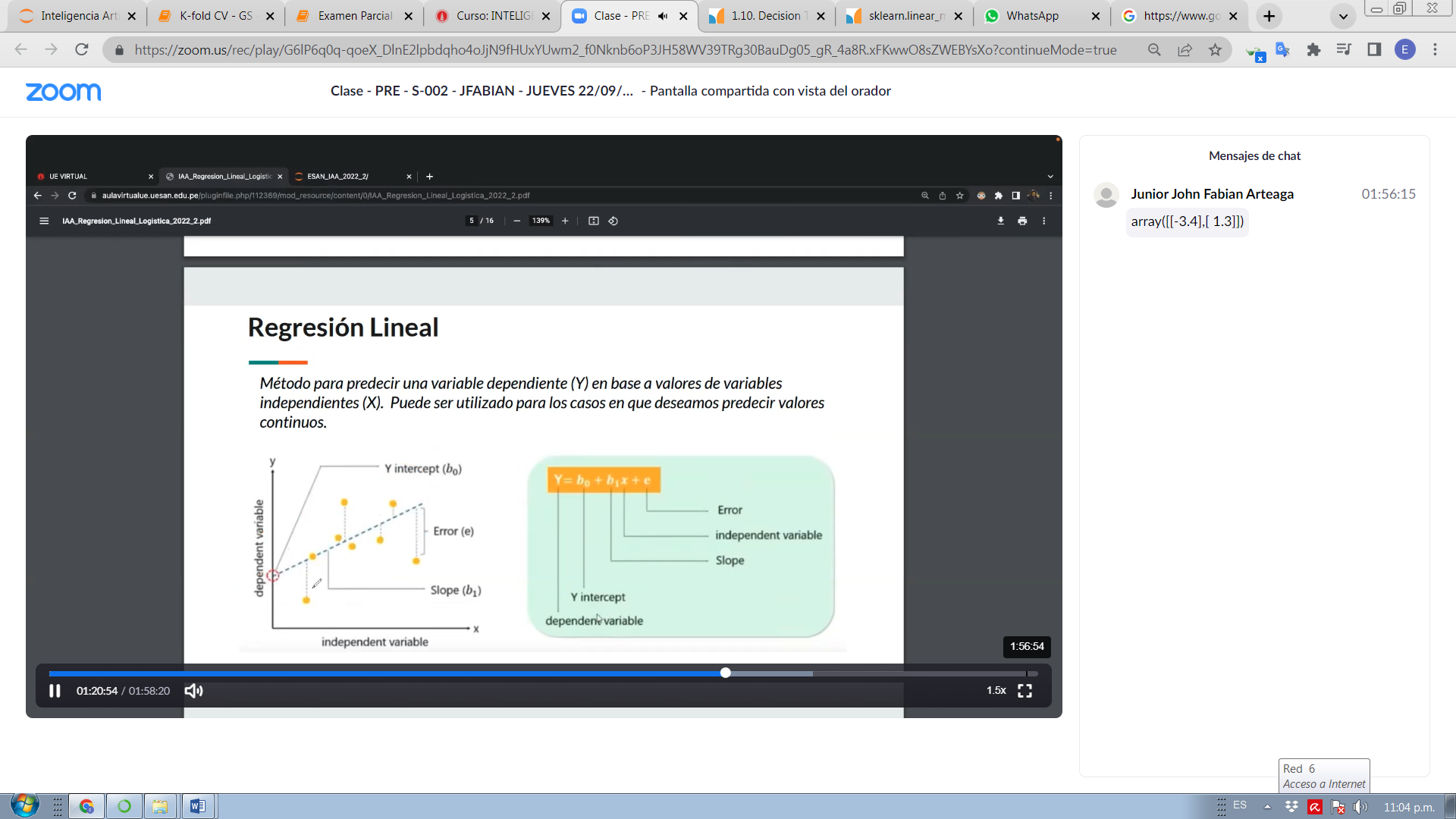
Inteligencia Artificial Avanzada 22 de septiembre de 2022

Internamente el Fit arma un árbol para ordenar la data

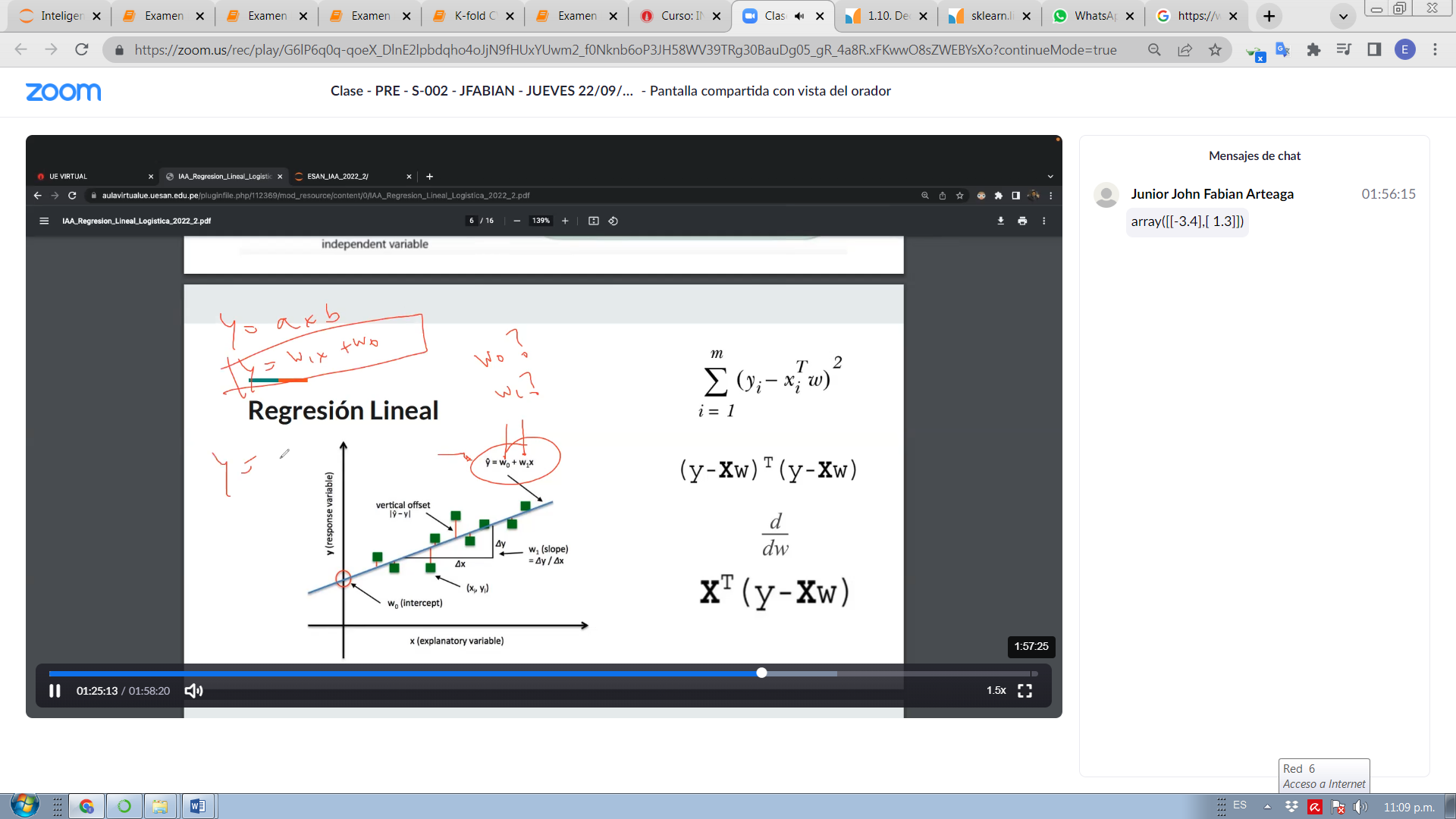
 

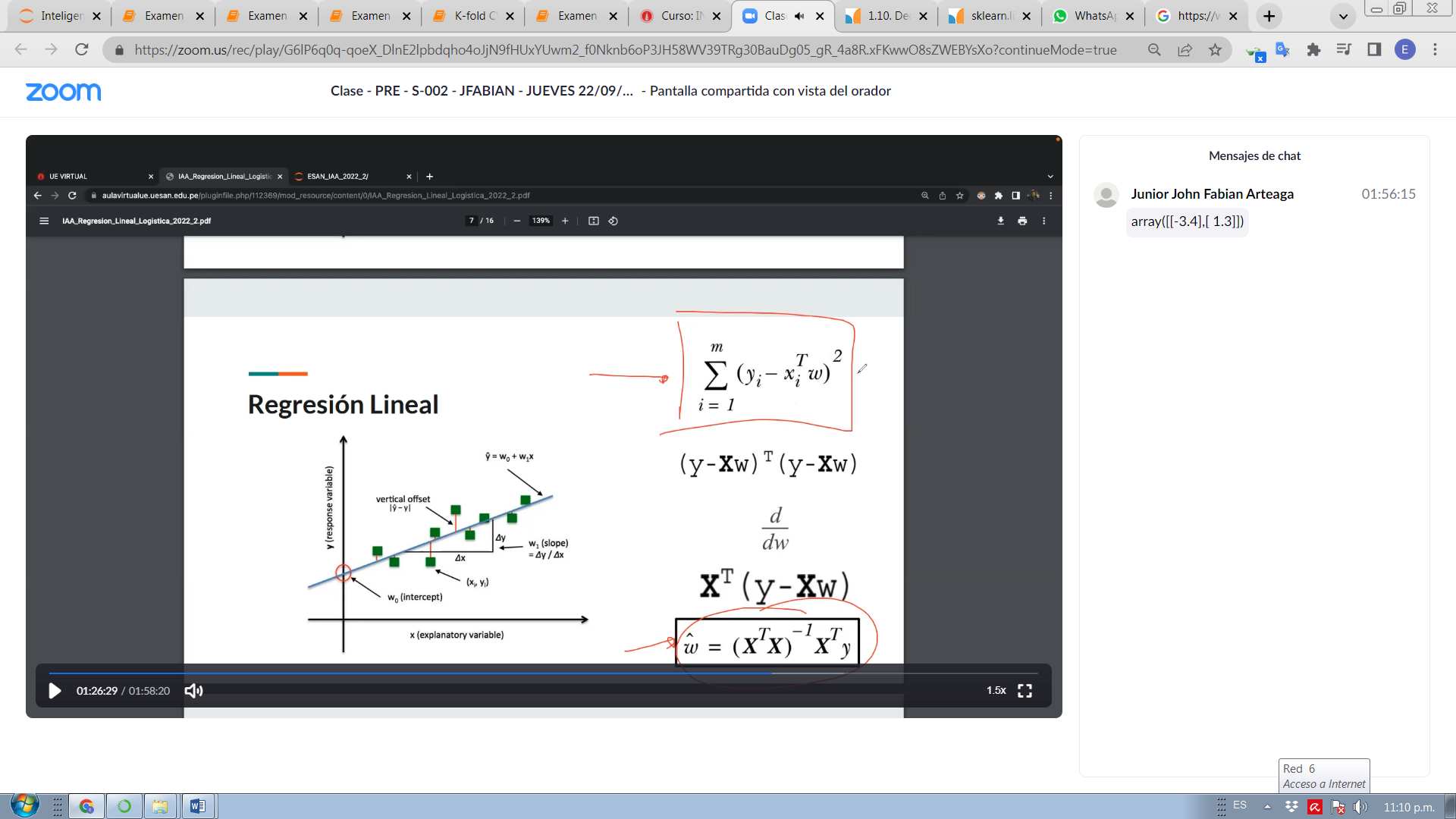
Para hallar esas distancias arma una especie de arbol donde sabe por ejemplo, que un nodo esta cerca de otro nodo. Comparar mediante arboles es más eficiente ya que genera lo que se conoce como busqueda ordenada. El KNN de Train se ordena antes de compararla con la data de Test.



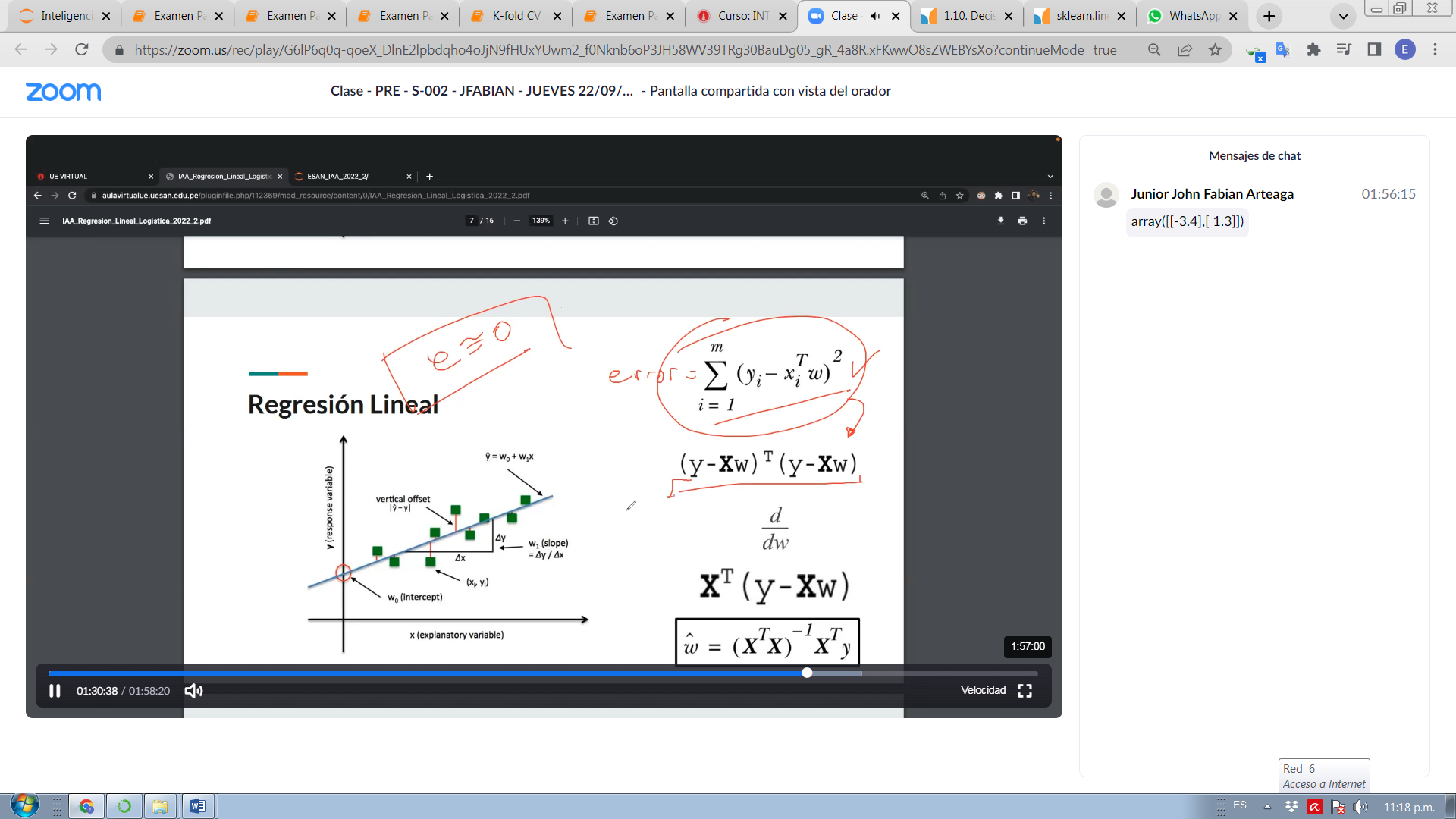
Por ejemplo si quieres lo 3 más cercanos ordenas tu data de menor a mayor y te quedas con los primeros. 

El que cause menos error de la distancia. A esto se le llama MSE o Media de los cuadrados de los errores.

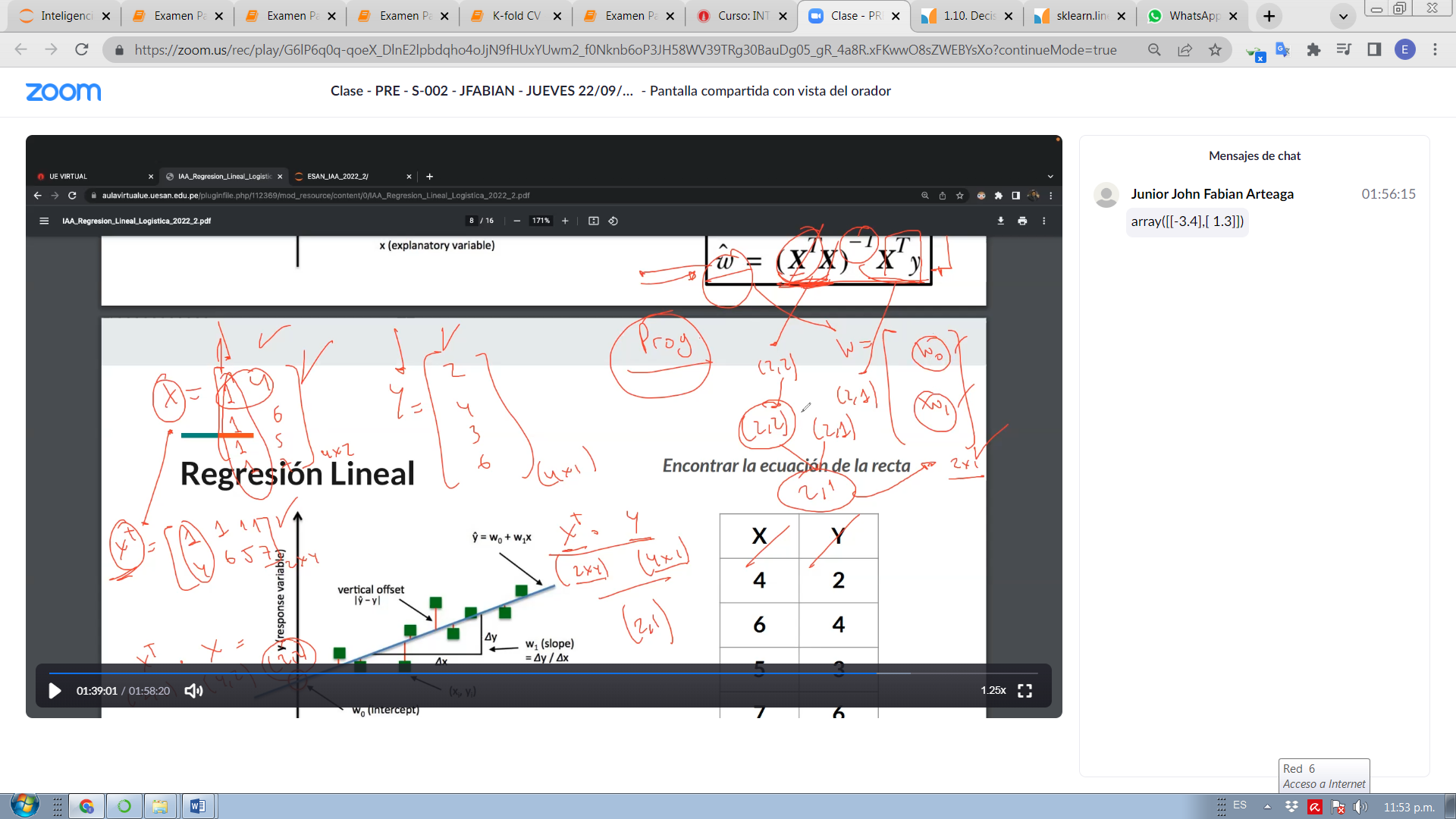




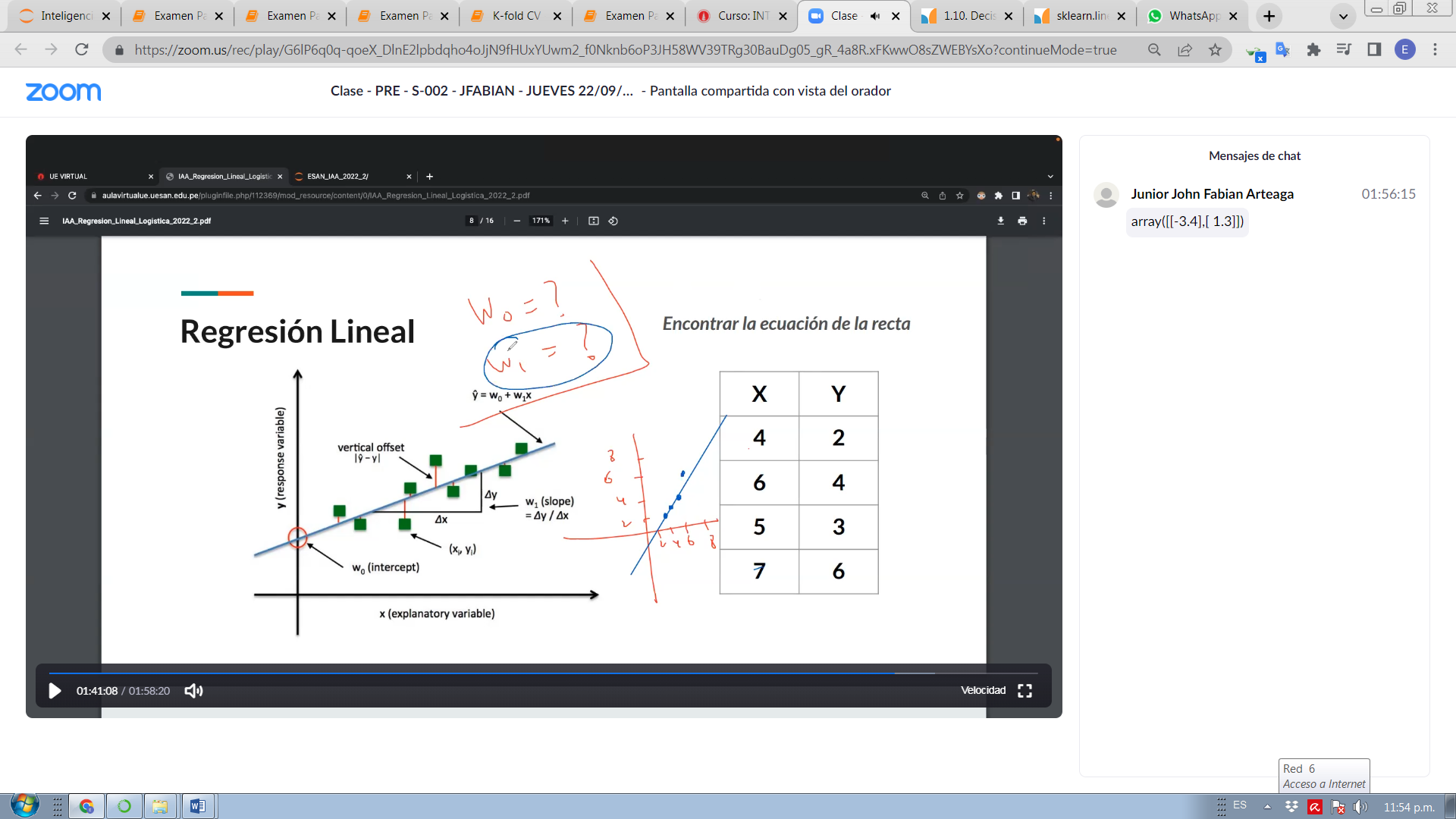
La fórmula que se muestra ahí arriba halla el error. La y es lo que yo conozco osea los puntos mientras xTw equivale a la ecuación de la recta de regresión. Es decir sale una sumatoria de distancias o suma de errores cuadrados. Obviamente se espera que el error sea cercano a 0.



La ecuación entre paréntesis y al medio, es la expresión en forma matricial. Ahora la derivada que esta abajo es para determinar los pesos para que sean mínimos. La formula que esta mas abajo sirve para hallar los pesos y esos pesos se supone que me dan la formula de la ecuación en la cual mi error es lo mas cercanos a 0.



El W0 es el intercepto y el otro es el coeficiente.



Pendiente positiva. W1 entonces es positivo



Wo se cruza con el eje Y en la parte negativa entonces es negativo.